

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 19, 2005, 23:00:56 ; Search time 651 Seconds

(without alignments)

2761.710 Million cell updates/sec

Title: US-10-659-004-104

Perfect score: 1599

Sequence: 1 MARKSTVILLTPALSTVNP.....FGIMLISIEBHKVPRINF 302

## Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 554816 seqs, 297661598 residues

Total number of hits satisfying chosen parameters: 11089632

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Command line parameters:

-MODEL=frame+ p2n.model -DEV=x1h  
-Q/cgn2\_1/USPTO.spool/US10659004/runat 14032005 181236 15793/app query.fasta\_1.455  
-DB=Published Applications NA -OPM=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPEC=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62  
-TRANSHUMAN40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pcr -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MARLEN=200000000 -USER=US10659004 @CGN 1 1 480 @runat 14032005 181236 15793  
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPLOCH=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAREXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

## Database :

Published Applications NA: \*  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq: \*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq: \*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq: \*  
4: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq: \*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq: \*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq: \*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq: \*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq: \*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq: \*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq: \*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq: \*  
12: /cgn2\_6/ptodata/2/pubpna/US09D\_NEW\_PUB.seq: \*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq: \*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq: \*  
15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq: \*  
16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq: \*  
17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq: \*  
18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq: \*  
19: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq: \*  
20: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq: \*  
21: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq: \*  
22: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1599	100.0	1101	19 US-10-659-004-103	Sequence 103, App
2	1359.5	85.0	2250	18 US-10-764-604-28	Sequence 28, App
3	1359.5	85.0	2744	15 US-10-171-581-170	Sequence 170, App
4	1359.5	85.0	2744	15 US-10-177-293-263	Sequence 263, App
5	1359.5	85.0	2744	17 US-10-172-118-1400	Sequence 1400, App
6	1359.5	85.0	2744	17 US-10-342-887-1400	Sequence 1400, App
7	1359.5	85.0	2744	18 US-10-769-612-1	Sequence 1, App
8	1359.5	85.0	3523	14 US-10-357-930-25736	Sequence 25736, A
9	1359.5	85.0	3523	14 US-10-198-846-10007	Sequence 10007, A
10	1356.5	84.8	2268	17 US-10-188-832-176	Sequence 176, App
11	1356.5	84.8	2268	17 US-10-659-004-105	Sequence 105, App
12	1356.5	84.8	3461	17 US-10-295-027-55	Sequence 55, App
13	1356.5	84.8	3461	17 US-10-295-027-838	Sequence 838, App
14	1356.5	84.8	3461	17 US-10-295-027-878	Sequence 878, App
15	1356.5	84.8	3461	17 US-10-058-270A-29	Sequence 29, App
16	1356.5	84.8	3609	19 US-10-855-588-21	Sequence 21, App
17	1355	84.7	1186	19 US-10-659-004-101	Sequence 101, App
18	1352.5	84.6	3445	17 US-10-099-322-25	Sequence 25, App
19	1352.5	84.6	3445	17 US-10-044-564-25	Sequence 97, App
20	1352.5	84.6	3445	19 US-10-659-004-97	Sequence 107, App
21	1351.5	84.5	2310	19 US-10-659-004-107	Sequence 3, App
22	1349.5	84.4	2776	17 US-10-182-033-3	Sequence 113, App
23	1347.5	84.3	3445	19 US-10-659-004-113	Sequence 111, App
24	1345.5	84.1	3445	19 US-10-659-004-111	Sequence 1, App
25	1324.5	82.8	2265	18 US-10-769-612-13	Sequence 13, App
26	1311.5	82.0	3461	17 US-10-182-033-1	Sequence 1, App
27	1311.5	82.0	3586	14 US-10-176-847-51	Sequence 51, App
28	1265.5	79.1	2211	19 US-10-659-004-109	Sequence 109, App
29	1007	63.0	2811	19 US-10-855-588-23	Sequence 23, App
30	875	54.7	1597	17 US-10-131-110-12	Sequence 12, App
31	831	52.0	551	17 US-10-182-033-12	Sequence 12, App
32	718	44.9	3309	18 US-10-357-930-23267	Sequence 23267, A
33	718	44.9	3309	18 US-10-357-930-29142	Sequence 29142, A
34	718	44.9	4417	10 US-09-814-353-22020	Sequence 22020, A
35	718	44.9	5231	16 US-10-269-909-39	Sequence 39, App
36	718	44.9	5231	16 US-10-723-860-4160	Sequence 4160, App
37	718	44.9	5330	9 US-09-789-561-14	Sequence 14, App
38	718	44.9	5330	19 US-10-883-936-14	Sequence 14, App
39	718	44.9	5446	18 US-10-322-696-5	Sequence 5, App
40	712	44.5	5299	18 US-10-723-860-8029	Sequence 8029, App
41	712	44.5	5748	18 US-10-322-696-2	Sequence 2, App
42	578	36.1	912	19 US-10-659-004-99	Sequence 99, App
43	514.5	32.2	366	10 US-09-918-995-30283	Sequence 30283, A
44	439	27.5	4573	9 US-09-880-107-1649	Sequence 1649, App
45	439	27.5	4573	10 US-09-960-706-489	Sequence 489, App

## ALIGNMENTS

RESULT 1  
US-10-659-004-103  
; Sequence 103, Application US/10659004  
; Publication No. US20050048507A1  
; GENERAL INFORMATION:  
; APPLICANT: Zhong et al.  
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD.  
; FILE REFERENCE: 21402-608  
; CURRENT APPLICATION NUMBER: US/10/659, 004  
; PRIOR APPLICATION NUMBER: 2003-09-09  
; PRIOR FILING DATE: 2001-06-04  
; PRIOR APPLICATION NUMBER: 60/295, 607  
; PRIOR FILING DATE: 2001-06-04  
; PRIOR APPLICATION NUMBER: 60/295, 661  
; PRIOR FILING DATE: 2001-06-04  
; PRIOR APPLICATION NUMBER: 60/296, 404  
; PRIOR FILING DATE: 2001-06-06  
; PRIOR APPLICATION NUMBER: 60/296, 418  
; PRIOR FILING DATE: 2001-06-06  
; PRIOR APPLICATION NUMBER: 60/297, 414  
; PRIOR FILING DATE: 2001-06-11



Qy 81 Aenllegiylleasplyleysargilleisilehisasaphisaphisese 100  
Db 241 AATTTAGCATAGATAGATTTAAAGATCCATATACCTAGACCACTACTCA 300  
Qy 101 Asphigluhihiiseseasphigluhihiiseseasphigluhihiisese 119  
Db 301 GACCAAGCATCTAGCATAGAGGCTGCTGAGCATGAGCATCTGAGACAC 360  
Qy 119 ----- 119  
Db 361 GAGATCACTGCAATATCATGCTGCTTGTGTAATAAGCGAAAGCTTTTGC 420  
Qy 119 ----- 119  
Db 421 CCAAGCATGACTAGATAGTTCAAGTAAGATCTTGAACAGCCAGGGAAGAGCT 480  
Qy 119 ----- 119  
Db 481 CACGACGAGAACATGCGAGGTGTAAGAGATGCAAGACAGTGTAGTGTAGAA 540  
Qy 119 ----- 119  
Db 541 GTGACTCACTGTGTACACACTGTCTGTAGAGAACTCTTGTAGAGCAATAGAG 600  
Qy 119 ----- 119  
Db 601 ACTCAAGACCTGGAACCTTCCCAAGATGAGACGCTCATCCACCAAGTGC 660  
Qy 119 ----- 119  
Db 661 ACATCAAGAGCCGGGTGAGCCGGCTGCTGTAGAAACAAATGATCTGTAGTAG 720  
Qy 119 ----- 119  
Db 721 CCCCCAAAAGCTTTATGTATTCAGAAACAAATGAAATCTCAGAGGTGTTCAAT 780  
Qy 119 ----- 119  
Db 781 GCATCAAGCTAGCATCTCATGAGCATGGGATCCAGGTTCCGGTGAATGCAAGAG 840  
Qy 119 ----- 119  
Db 841 TTCAACTATCTCTGCCAGCATCATCAACCAATGTAGTGTATCTGTGATTCAT 900  
Qy 119 ----- 119  
Db 901 ACAATGAAAGAGGCTGAATCCCTCAAGACCTATTCATTACAAATAGCCGGGT 960  
Qy 119 ----- 119  
Db 961 GGTGTTTTATAGCCATTTCATCATCATGTTCTGTCTGTGGGGTATCTTAGTG 1020  
Qy 119 ----- 119  
Db 1021 CCTCTCATGAATCGGGGTGTTTTCAAAATTTCTCTGAGTTCTTGTGCGACGTGCGCTT 1080  
Qy 119 ----- 119  
Db 1081 GGGACTTGAATGAGATGTTTTTTTACACCTTCTTCACATTCATGCAAGTACACAC 1140  
Qy 119 ----- 119  
Db 1141 CATATGATAGCCATGAGAAACAGCAATGAAATGAAAGAGACCACTTTTCAATCAT 1200  
Qy 119 ----- 119  
Db 1201 CTGTCTTCTCAAAACATGAGAAAGTCCCTATTTTGTATTCACGCTGAGAGGCTTACA 1260  
Qy 119 ----- 119  
Db 1261 GCTTAGAGAGCCGTATTTTCATGTTTCTGTGTGAACATGCTTCATGATGAAACAA 1320  
Qy 119 ----- 119

Db 1321 TTTAAGATAGAGAAAGAAATCAGAGAAACCTGAAATGATGATGTGAGATT 1380  
Qy 119 ----- 119  
Db 1381 AAGAGAGTGTTCAGATGATGATCTCACTTTCAACAAATGAGAGAAAGTATACA 1440  
Qy 119 ----- 119  
Db 1441 GATGATGAATGAGGCTATTTACAGACAGACTCACAAGAGCCCTCCACTTGTATTCT 1500  
Qy 119 ----- 119  
Db 1501 CAGAGCCTGCACTTGTGAAGAAAGAGGTATGATGCTATGCTATTCACAGSAA 1560  
Qy 119 ----- 119  
Db 1561 GTCTACATGATATGTATACCCAGAGGTGCAGAAATTAATGCCATTCATTTCCAGAT 1620  
Qy 119 ----- 119  
Db 1621 ACACGCGCAGTCAGAGATCTCATTCACCATCATGACTATCATATTTCTCCAT 1680  
Qy 120 ----- 119  
Db 1681 CATCAACCAACCAAAACCACTCTCAGAGTACAGCCAGCCCTACTCTCGGAGAGAG 1740  
Qy 134 LeulysapalaglyValAlaThrLeuAlaTrpMetValIleMetGlyAspGlyLeuHis 153  
Db 1741 CTGAAGATGCGCGCTGCGCACCTTGTGCTGATGTGATTAATGGGATGAGGCTCGAC 1800  
Qy 154 AsnheserapglyleuAlaIleGlyAlaIlePheThrGluGlyLeuSerSerGlyLeu 173  
Db 1801 AATTCAGCGATGCGCTTACGATGAGTGTGCTTTTACGAAAGCTTATTCAGTGTGTTA 1860  
Qy 174 SerThrservalAlaValAlaPheCysHisGluLeuProHisGluLeuGlyAspPheAlaVal 193  
Db 1861 AGTACTTCTGTGCTGTGCTGTGTCTGATGATGAGTGTGCTGCTCATGATTAAGTGTGCTGCT 1920  
Qy 194 LeuleuysalaglyMetThrValIleGlnAlaValLeuTyraAlaLeuSerAlaMet 213  
Db 1921 CTACTAAAGGCTGACATGACCGTTAAGCAGGCTGCTTATTAATGATGTGCGCATG 1980  
Qy 214 LeuAlaTyrlleuGlyMetAlaThrGlyIlePheIleGlyHisTyraIleGluAsnValSer 233  
Db 1981 CTGGCGTATCTTGAAGTGGCAACAGAAATTTCAATGTGCTATTAATGCTGAAGAAATGTTCT 2040  
Qy 234 MetTrpIlePheAlaLeuThrAlaGlyLeuPheMetHisValAlaLeuValAspMetVal 253  
Db 2041 ATGTGAAATTTGCACTTACTGTGCTGTGCTTATTCATGATGTGCTGTGATGATGTGTA 2100  
Qy 254 ProGluMetLeuHisAsnAspAlaSerAspHisGlyCysSerArgTrpGlyTyrlPhePhe 273  
Db 2101 CTTGAATGCTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160  
Qy 274 LeuGlnAsnAlaGlyMetLeuLeuGlyPheGlyIleMetLeuLeuIleSerIlePheGlu 293  
Db 2161 TTACAGAAATGCTGGAGATGCTTTTGGGTTTGAATTAATGTTACTTAATTCATTTGAA 2220  
Qy 294 HisIleValIlePheArgIleAsnPhe 302  
Db 2221 CATAAATGCTGTTCGATTAATAATTTT 2247

RESULT 3  
US-10-171-581-170  
; Sequence 170, Application US/10171581  
; Publication No. US20030104426A1  
; GENERAL INFORMATION:  
; APPLICANT: Dai, Hongyue  
; APPLICANT: Linsley, Peter  
; TITLE OF INVENTION: Signature Genes in Chronic Myelogenous Leukemia  
; FILE REFERENCE: 9301-157-999

CURRENT APPLICATION NUMBER: US/10/171,581  
CURRENT FILING DATE: 2002-06-14  
PRIORITY APPLICATION NUMBER: 60/298,914  
PRIORITY FILING DATE: 2001-06-18  
NUMBER OF SEQ ID NOS: 366  
SEQ ID NO 170  
LENGTH: 2744  
TYPE: DNA  
ORGANISM: Homo sapiens  
PUBLICATION INFORMATION:  
DATABASE ACCESSION NUMBER: U41060  
DATABASE ENTRY DATE: 2001-06-18  
US-10-171-581-170

Alignment Scores:  
Pred. No.: 1,31e-150 Length: 2744  
Score: 1359.50 Matches: 301  
Percent Similarity: 40.32% Conservative: 1  
Best Local Similarity: 40.19% Mismatches: 0  
Query Match: 85.02% Indels: 447  
DB: 15 Gaps: 1

US-10-659-004-104 (1-302) x US-10-171-581-170 (1-2744)

QY 1 MetAlaArgLYbLeuSerValIleuThrPheAlaLeuSerValThrAsnPro 20  
DB 138 ATGGGAGAGAGTATCTGTAACTTGAATCCGACCTTGCCCTCTGTCAAAATCCC 197

QY 21 LeuHISGLuLeuLYaAlaAlaIlePheProGlnThrGluLYIleSerProAsnTrp 40  
DB 198 CTTGATGAACCTAAAGCAGCTCTTCCCGCAGACCATGAGAAATTAAGTCGAAATGG 257

QY 41 GLuSerGlyIleAsnValAspLeuAlaIleSerThrArgGlnTrpHISLeuGlnIleu 60  
DB 258 GAATCTGCATTAATGTTGACTTGGCAATTCACACGCAATATCATCTACACAGCTT 317

QY 61 PheTYArgTYRGLyGluAsnAsnSerLeuSerValGluLYPheArgLYbLeuGln 80  
DB 318 TTCTACCGCTATGAGAAATATATCTTTGTGAGTTGAAAGGTTGAGAAATTAATCTCA 377

QY 81 AsnIleGlyIleAspLYbLeuLYaArgIleHISleHISAspHISAspHISAspHISer 100  
DB 378 AATATAGCATGATGATTAAGTTAAAGATTCATATACCATGACACGACCATCATCTCA 437

QY 101 AspHISGLuHISISerAspHISGLuArgHISerAspHISGLuHISerAspHISer 119  
DB 438 GACCAAGCATCACTGACAGCATGAGCGTCACTCAGACCATGACCATCATCTCAGACAC 497

QY 119 ----- 119

DB 498 GAGCATCACTGACCATATATCATGCTGCTTGTGTAATAATAGCGAAAGCTTTTGC 557

QY 119 ----- 119

DB 558 CCAGACATGACTAGATAGTTCAGGTAAAGATCTAGAAACAGCCAGGGAAAGAGCT 617

QY 119 ----- 119

DB 618 CACCGACGAGACATGCGCAGGTGAGAAAGATGTCAGAGCAGTGTAGTGTAGTAA 677

QY 119 ----- 119

DB 678 GTGACCTCAACTGTGTAAACAACACTGTCTGTGAAGAACTCACTTTCTAGACAAATAG 737

QY 119 ----- 119

DB 728 ACTCCAGAGCTGAAAACCTTTCCCAAAAGATGTAAGCAGCTCCACTCCACCCAGTGC 797

QY 119 ----- 119

DB 798 ACATCAAGAGCGGGGTAGCGGCTGCTGTAGAGAAACAAATGAATCTGTGAGTGA 857

QY 119 ----- 119

DB 858 CCCCCAAAAGCTTTATGATATCCAGAAACAAATGAAAATCCTCAGAGTGTTCAT 917

QY 119 ----- 119

DB 918 GCATCAAGCTACTGACATCTCATGCGATGGCATCCAGGTTCCGTGAATGCAACAGAG 977

QY 119 ----- 119

DB 978 TTCAACTATCTGTCCAGCCATCATCAACCAATTGATGCTAGATCTGTCTGATTCAT 1037

QY 119 ----- 119

DB 1038 ACAAGTGAAGAGCTGAATTCCTCCAAAGCCTATTCTATACAAATAGCCTGGGT 1097

QY 119 ----- 119

DB 1098 GGTGTTTATAGCATTTCCATTCATGATTCCTGTCTGTGCGGGTTATCTTAGTG 1157

QY 119 ----- 119

DB 1158 CCTCTCATGAATCGGGTGTTCCTCAATTTCTCTGAGTTCTTGTGGCACTGGCGGTT 1217

QY 119 ----- 119

DB 1218 GCGACTTGAGTGTGATGCTTTTACACCTTCTTCCACATTCATGCAAGTCACAC 1277

QY 119 ----- 119

DB 1278 CATATGATAGCATGAAGAACACGATGGAATGAAGAGACCACTTTTCAGTCAT 1337

QY 119 ----- 119

DB 1338 CTGTCTTCTCAAAACATGAAGAAAGTCCATTTTGATTCACAGTGAAGGTTCTACA 1397

QY 119 ----- 119

DB 1398 GCTTAGAGGCGCTGATTTTCATGTTCTTGTGAACATGTCCTCACATGATCAACAA 1457

QY 119 ----- 119

DB 1458 TTTAAGATGAAGAAAGAAAGATCAGAGAACTGAAATGATGATGTGAGATT 1517

QY 119 ----- 119

DB 1518 AAGAGCAGTTGTCAGATGAATTCACCTTTCAACAAATGAGAGAAAGTATACA 1577

QY 119 ----- 119

DB 1578 GATGATGAAGTGAAGGCTATTTACAGACAGACTCACAAAGGCCCTCCACTTGTATCT 1637

QY 119 ----- 119

DB 1638 CAGCAGCTGCACTTTGGAAGAAAGAGTCAATGATGCTATGCTATCCACAGAA 1697

QY 119 ----- 119

DB 1698 GTCTAATGAATATGTAACCCAGAGGTGCAAGATAAATGCCATTCACTTCCAGAT 1757

QY 119 ----- 119

DB 1758 ACACCTGGCAGTGAAGCATCTCATTCACACCATCATGATCATATATCTTCAT 1817

QY 120 ----- 133

DB 1818 CATCACACCAACCAACCATCTCAGAGTCAAGCCAGCCCTACTCTTCGGAGAGAG 1877

QY 134 LeuLYAspAlaGLyValAlaThrLeuAlaTrpMetValIleMetGLyAspGLyLeuHIS 153

DB 1878 CTGAAGATGCGCGCTGCGCACTTGGCTTGATGATGAATGGGTGATGGCTGCAC 1937

QY 154 AsnPheSerAspGLyLeuAlaIleGLyAlaAlaIlePheThrGluGlyLeuSerSerGLyLeu 173

```

Db 1938 AATTTCAGCAGATGGCTAGCAATGTGCTGCTTTACTGAAAGCTTATCAAGTGTTA 1997
Qy 174 SerThrSerValAlaValPheCysHisGluLeuProHisGluLeuGlyAspPheAlaVal 193
Db 1998 AGTACTTCTGTGCTGCTGCTTCTGTCATGATGCTGCTCAATGATGATGATGCTGCTGTT 2057
Qy 194 LeuLeuValAlaGlyMetThrValLysGlnAlaValLeuTyrAsnAlaLeuSerAlaMet 213
Db 2058 CTACTAAAGCTGCGCATACCTTAAAGAGCTGCTCTTATATGATTCATTCAGCCCATG 2117
Qy 214 LeuAlaTyrLeuGlyMetAlaThrGlyLysPheIleGlyHisTyrAlaGluAsnValSer 233
Db 2118 CTGGCGTATCTTGAATGCAACAGAAATTTTCATTGTCATTATGCTGAAATGCTTCT 2177
Qy 234 MetTrpIlePheAlaLeuThrAlaGlyLeuPheMetHisValAlaLeuValAspMetVal 253
Db 2178 ATGCGAATATTGCACTTACTGCTGCTGCTTATTCATGATGATGCTGCTGCTGATGCTGA 2237
Qy 254 ProGluMetLeuHisAsnAspAlaSerAspHisGlyCysSerArgTyrGlyTyrPhePhe 273
Db 2238 CTGAAATGCTGCACAAATGATGCTAGTACCATGATGATGATGATGATGATGATGATGAT 2297
Qy 274 LeuGlnAsnAlaGlyMetLeuLeuGlyPheGlyLysMetLeuLeuIleSerIlePheGlu 293
Db 2298 TTACGATGCTGCGGATGCTTTGGGTTTGGAAATTAATTAATTAATTAATTAATTAATTA 2357
Qy 294 HisLysIleValPheArgIleAsnPhe 302
Db 2358 CATAAATCGTGTTCGTATAAATTC 2384

RESULT 4
US-10-177-293-263
; Sequence 263, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Ganavarpu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongzao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Baet Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Puzstai, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
; TITLE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; PRIOR APPLICATION NUMBER: 2002-06-21
; PRIOR FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 263

```

```

; LENGTH: 2744
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-177-293-263

Alignment Scores:
Pred. No.: 1,31e-150 Length: 2744
Score: 1359.50 Matches: 301
Percent Similarity: 40.32% Conservative: 1
Best Local Similarity: 40.19% Mismatches: 0
Query Match: 85.02% Indels: 447
DB: 15 Gaps: 1

US-10-659-004-104 (1-302) x US-10-177-293-263 (1-2744)
Qy 1 MetAlaArgLysLeuSerValIleLeuIleLeuThrPheAlaLeuSerValThrAsnPro 20
Db 138 ATGGCGAAGATATATCTGTATCTTGAATCCTGATCCTTGGCCCTCTGTCACAAATCCC 197
Qy 21 LeuHisGluLeuValAlaAlaPheProGlnThrThrGlyLysIleSerProAsnTrp 40
Db 198 CTTCAATGAATPAAAGCAGCTGCTTCCCAAGACACTGAGAAATTAATGTCGAATTGG 257
Qy 41 GluSerGlyIleAsnValAspLeuAlaIleSerThrArgGlnTyrHisLeuGlnGlnLeu 60
Db 258 GAATCTGCAATTAATGTTGACTTGCACTTGCAATTTCCACACGCAATATCATCTACAAAGCTT 317
Qy 61 PheTyrArgTyrGlyGluAsnAsnSerLeuSerValGluGlyPheArgLysLeuGln 80
Db 318 TTCTACCGCTATGAGAAATATATCTTGTTCAGTTGAAGGCTTCAGAAATTAATCTTCAA 377
Qy 81 AsnIleGlyIleAspLysIleLeuArgIleHisIleHisIleAspHisIleSer 100
Db 378 AATATAGCATPAAATGATPAAAGATPAAAGATPAAAGATPAAAGATPAAAGATPAAAGAT 437
Qy 101 AspHisGluHisIleSerAspHisGluLysHisIleSerAspHisGluHisIleSerAsp 119
Db 438 GACCAAGCATCACTCACTGACCAATGAGCGCTCACTCAAGCATCACTCAAGCATCACT 497
Qy 119 ----- 119
Db 498 GAGCATCACTGACCAATATCATGCTGCTCTGCTGTAATAAATAAGCAAAAGCTTTGCG 557
Qy 119 ----- 119
Db 558 CCAGACATGACTCAGATAGTTCAAGTAAAGATCTTAGAAACAGCCAGGGAAAGAGCT 617
Qy 119 ----- 119
Db 618 CACCGACCAAGACATGCCAGTGTGAAGAAAGATGCAAGACAGTGTAGTGTGA 677
Qy 119 ----- 119
Db 678 GTGACCTGAATGTGTACAACACTGTCTGTGAAGAACTCACTTCTAGAGCATAGAG 737
Qy 119 ----- 119
Db 738 ACTCAAGACTGGAAATCTTCCCAAAAGATGTAAGACAGCTCCATCCACCAAGTGC 797
Qy 119 ----- 119
Db 798 ACATCAAAAGCCGGGTGAGCCGCTGCTGTGTAGAAAACAATAATTAATCTGTAGTAG 857
Qy 119 ----- 119
Db 858 CCCGAAAGGCTTTATGTATTTCCAGAAACACAATGAATTAATCTCAGAGAGTTCAT 917
Qy 119 ----- 119
Db 918 GCATCAAGACTAGACATCTCAGTGCATGGGATCCAGGTTCCGCTGATGACACAGAG 977
Qy 119 ----- 119

```

```

Db 978 TTCACTATCTCTGTCAGCCATCATCAACCAATGTGCTGATCTTGTCTGATTCAT 1037
QY 119 ----- 119
Db 1038 ACAAGTAAAAAGGCTGAAATCCCTCCAAAGCTTATTCATAAATAGCCTGGGTT 1097
QY 119 ----- 119
Db 1098 GGTGTTTTATAGCATTTCATCATCAATGTTCTCTGCTGCTGGGGTATATCTTAGTG 1157
QY 119 ----- 119
Db 1158 COTCATGAATCGGGTGTGTTTTCAATTTCTCTGAGTTTCTTGCGACATGCGCGGTT 1217
QY 119 ----- 119
Db 1218 GGGACTTGTAGTGTGATGCTTTTTTACACCTTCTTCACATTCATGCAAGTCACGAC 1277
QY 119 ----- 119
Db 1278 CATAGTCAATGCGATGAAGAACCCAGCAATGCAATGAAAAAGAGACCACTTTTCAGTCAT 1337
QY 119 ----- 119
Db 1338 CTGCTCTCTCAAAACATGAAGAAGTGCCTATTTTGTATTCACGTCGAGGGTCTACAA 1397
QY 119 ----- 119
Db 1398 GCTCTAGGAGGCTGTATTTTCATGTTCTGTGTGAACATGTCCTCATTTGATCAACAA 1457
QY 119 ----- 119
Db 1458 TTTAAAGATAGAGAAAAAAGATCAGAAAGAACTGAAAAATGATGATGATGTGAGATT 1517
QY 119 ----- 119
Db 1518 AAGAAGAGTGTGTCCAGATGATCATCTCACTTTCAACAAATGAGAGAAAGATGATACA 1577
QY 119 ----- 119
Db 1578 GATGATCGAAGTGAAGGCTATTTTACGAGACATCAAGAGCCCTCCCATTTGATTCCT 1637
QY 119 ----- 119
Db 1638 CAGCAGCTGCACTGTGGAAGAAAGAGTCAATGATGATGCTCATGCTCACAAGAA 1697
QY 119 ----- 119
Db 1698 GTCTCAATGAATATGTATCCAGAGGTCAGAAATAAATGCAATTCACATTTCCAGCAT 1757
QY 119 ----- 119
Db 1758 ACACCTGGCGAGCTCAGACGATCTCATTCACACCATCATGACTACATCATATTTCTCAT 1817
QY 120 ----- 113
Db 1818 CATCACACACCAACCAACCACTCTCACTCAGACGACGCTACTCTCTGGAGAGAG 1877
QY 134 LeuLyAspAlaGlyValAlaThrLeuAlaTrpMetValIleMetGlyAspGlyLeuHis 153
Db 1878 CTGAAGAAGATGCGCGCTGCGCCACTTTGGCTGATGTGATATATGGGTGATGGCTGCAC 1937
QY 134 AenPheSerAspGlyLeuAlaIleGlyValAlaAlaPheThrGluGlyLeuSerSerGlyLeu 173
Db 1938 AATTTCAGGATGCGCTTACCAATGTGTCTTTTACGAAGGCTTAACTCAAGTGTATTA 1997
QY 174 SerThrSerValAlaValPheCysHisGluLeuProHisGluLeuGlyAspPheAlaVal 193
Db 1998 ACTTACTTCTGCTGTGTGTGTCTCATGAGTTGCTCCATGAATTAAGTGACCTTTGCTGTT 2057
QY 194 LeuLeuValAlaGlyMetThrValIlyGlnAlaValLeuTyrAsnAlaLeuSerAlaMet 213
Db 2058 CTACTAAGGCTGGCATGACCGTTAAGCAGGCTGTCTTTATATGATGATGTTCAGGCATG 2117

```

```

QY 214 LeuAlaTyrLeuGlyMetAlaThrGlyIlePheIleGlyHisTyrAlaGluAsnValSer 233
Db 2118 CTGGCGATCTCTGGAAATGGCAACGAAATTTTCTTGTGCTCATTTGCTGAAAAATGTTCT 2177
QY 234 MetTyrIlePheAlaLeuThrAlaGlyLeuPheMetHisValAlaLeuValAspMetVal 253
Db 2178 ATGTGATATTTGCACTTACTGCTGCTGCTTATTCATGATGTGCTGTGATATATGCTA 2237
QY 254 ProGluMetLeuHisAsnAspAlaSerAspHisGlyCysSerArgTyrPhePhe 273
Db 2238 CTTGAATGCTGCACCAATGATGCTAGTACCATGATGATGACCCCTGGGGTATTTCTT 2297
QY 274 LeuGlnAsnAlaGlyMetLeuGlyPheGlyIleMetLeuLeuIleSerIlePheGly 293
Db 2298 TTACAGAAATGCTGGAATGCTTTTGGCTTTTGGATTTATGTTACTTATTTCCATTTTGA 2357
QY 294 HisIlyIleValPheArgIleAsnPhe 302
Db 2358 CATAAATCGTGTTCGATATAATTTTC 2384

RESULT 5
US-10-172-118-1400
; Sequence 1400, Application US/10172118
; Publication No. US2003024374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Robert, Chris
; APPLICANT: Van de Vijver, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1400
; LENGTH: 2744
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_012319
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1400

Alignment Scores:
Pred. No.: 1,336-150 Length: 2744
Score: 1359.50 Matches: 301
Percent Similarity: 40.32% Conserved: 1
Best Local Similarity: 40.19% Mismatches: 0
Query Match: 85.02% Indels: 447
DB: 17 Gaps: 1

US-10-659-004-104 (1-302) x US-10-172-118-1400 (1-2744)
QY 1 MetAlaArgIlyLeuSerValIleLeuIleLeuThrPheAlaLeuSerValThrAsnPro 20
Db 138 ATGGGAGAGAGATTAATCTGATCTGATCTGACCTTGGCCCTCTGTGCAAAATCCC 197
QY 21 LeuHisGluLeuValAlaAlaPheProGlnThrThrGluValIleSerProAsnTyrP 40
Db 198 CTTATGAATCAAAAGCAGCTGCTTTCCCGACACCACTGAGAAATTAATGTCGGAATGG 257
QY 41 GluSerGlyIleAsnValAspLeuAlaIleSerThrArgGlnTyrHisLeuGlnGlnLeu 60
Db 258 GAATGTGCAATTAATGTTGACTTGGCAATTTTCACAGGCAATATCATCTACACAGCTT 317
QY 61 PheTyrArgTyrGlyGluAsnAsnSerLeuSerValIleGluGlyPheArgIlyLeuGln 80

```



APPLICANT: Lindey, Peter S.  
APPLICANT: Mac, Mao  
APPLICANT: Roberts, Christopher J.  
APPLICANT: Van 't Veer, Laura Johanna  
APPLICANT: Van de Vijver, Marc J.  
APPLICANT: Bernards, Rene  
TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients  
FILE REFERENCE: 9301-188-999  
CURRENT APPLICATION NUMBER: US/10/342,887  
CURRENT FILING DATE: 2003-01-15  
PRIOR APPLICATION NUMBER: 60/298,918  
PRIOR FILING DATE: 2001-06-18  
PRIOR APPLICATION NUMBER: 60/380,710  
PRIOR FILING DATE: 2002-05-14  
PRIOR APPLICATION NUMBER: 10/172,118  
PRIOR FILING DATE: 2002-06-14  
NUMBER OF SEQ ID NOS: 2699  
SEQ ID NO 1400  
LENGTH: 2744  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-342-887-1400

Alignment Scores:  
Pred. No.: 1,316-150 Length: 2744  
Score: 1359.50 Matches: 301  
Percent Similarity: 40.32% Conservative: 1  
Best Local Similarity: 40.19% Mismatches: 0  
Query Match: 85.02% Indels: 447  
DB: 17 Gaps: 1

US-10-659-004-104 (1-302) x US-10-342-887-1400 (1-2744)

QY 1 MetAlaArgLysLeuSerValIleLeuIleLeuThrPheAlaLeuSerValThrAsnPro 20  
DB 138 ATGGGAGAGAGTATCTGTATCTTGATCTTGACCTTGACCTTGACCTTGACCTTGACCT 197  
QY 21 LeuHISGLIuLeuLysAlaAlaIlePheProGlnThrGluLysLeuSerProAsnTrp 40  
DB 198 CTTCATGAACTTAAAGCAGCTGCTTCCCGACACCATGAGAAATTAAGTCCGAAATGG 257  
QY 41 GluSerGlyIleAsnValAspLeuAlaIleSerThrArgGlnTyrHisLeuGlnGlnLeu 60  
DB 258 GAATCTGGATTAATGTTGACTTGGCAATTTCCACGCGCATATCATCTACAAACAGCTT 317  
QY 61 PheTyrArgTyrGlyGluAsnAsnSerLeuSerValGluGlyPheArgLysLeuLeuGln 80  
DB 318 TTCTACCGCTATGAGAAATTAATCTTGTCTGATGAAAGGCTTCAGAAATTAATCTTCAA 377  
QY 81 AsnIleGlyIleAspLysIleLysArgIleHisIleHisAspHisAspHisSer 100  
DB 378 AATATATGATATGATTAAGTATTAAGATTCATATACCATATGACATGACGACCATCTCA 437  
QY 101 AspHisGluHisHisSerAspHisGluArgHisSerAspHisGluHisHisSerAsp 119  
DB 438 GACCAACGACATCATCATGACATGAGCTCACTCAGACCATGACATCATCATGACCAAC 497  
QY 119 ----- 119  
DB 498 GAGCATCACTTGACCATATCATGCTGCTTGTGTAATAATATGAGCAAAAGCTTTTGC 557  
QY 119 ----- 119  
DB 558 CCAAGCATGACTCATATGTTCTCAGTAAAGATCTTCAAAACAGCAGGAGGAGAGAGCT 617  
QY 119 ----- 119  
DB 618 CACCGACCAAGACATGCCAGTGTAGAGAAATGTCAGAGACAGTGTAGTCTAGTAA 677  
QY 119 ----- 119  
DB 678 GTGACCTCAACTGTGTACAACACTGTCTCTGAGGAACACTCTTTCTAGAGCAATATAG 737

QY 119 ----- 119  
DB 738 ACTCGAAGACTGGAATACTTTCGCCAAGATGTAAAGACGTCCATCCACCAAGTCTC 797  
QY 119 ----- 119  
DB 798 ACATCAAGAGCCGGGTGAGCCGGCTGGTGTAGAAAACAATGAATCTGTGTAGTGTAG 857  
QY 119 ----- 119  
DB 858 CCGGAAAAAGCTTTATGTATTCAGAAACAAATGAAAAATCCTCAGAGGTGTTCAAT 917  
QY 119 ----- 119  
DB 918 GCATCAAGACTATGACATCATGCGATGCGGATCCAGGTTCCGGTGAATGCAACAGAG 977  
QY 119 ----- 119  
DB 978 TTCAACTATCTGTGCCAGCATCATCAACCAAAATGATGCTAGATCTTGTGTATTCAT 1037  
QY 119 ----- 119  
DB 1038 ACAAGTAAAGAAAGCTGAATTCCTCCAAAGCTTATTCATTACAAATAGCTGGGTT 1097  
QY 119 ----- 119  
DB 1098 GGTGTTTATAGCCATTTCATCATCATGATTCCTGTCTGTCTGTGGGTTATCTTAGTG 1157  
QY 119 ----- 119  
DB 1158 CCTCTCATGATCGGGTGTGTTTTCAAATTTCTCCTGAGTTTCCTGTGGCACTGGCCGTT 1217  
QY 119 ----- 119  
DB 1218 GGGACTTGAAGTGATGCTTTTTCACCTTTCACATTCATCATGCAATGACACAC 1277  
QY 119 ----- 119  
DB 1278 CATAGCTATAGCATGAAGAACCAAGCATGGAATGAAGAAAGACCACTTTCAGTAT 1337  
QY 119 ----- 119  
DB 1338 CTGTCTTCTCAAAACATGAAAGAAAGTCCATTTTGTATTCACGTGGAAGGTTTACA 1397  
QY 119 ----- 119  
DB 1398 GCTTAGAGGCTGTATTTTCATGTTCTTGTGTGAACATGTCCTCATATTGATCAACAA 1457  
QY 119 ----- 119  
DB 1458 TTTAAAGTAAAGAAAGAAAGATCAGAAAGAACTGAAATGATGATGTGGAGATT 1517  
QY 119 ----- 119  
DB 1518 AAGAGCAGTTGTCCAGATATGAATCTCACTTTCAACAAATGAGAGAAAGATATCA 1577  
QY 119 ----- 119  
DB 1578 GATGATGAACTGAAGCTATTTTACAGACAGATCACAAGACCTCCCATCTTGATTTCT 1637  
QY 119 ----- 119  
DB 1638 CAGGACCTGCAAGTCTTGAAAGAAAGAGGTCATGATAGCTCATGCTCATCCACAGAA 1697  
QY 119 ----- 119  
DB 1698 GTTACATGAATATGTACCAGAGGTGCAAGATAAATGCCATTCATTTCCACGAT 1757  
QY 119 ----- 119  
DB 1758 AACTGGCCAGTACAGACATCTCATTCACCAATCATATGACTATCATATTTCTCAT 1817  
QY 120 ----- 120  
DB ----- HisHisProHisSerHisSerGlnArgTyrSerArgGlu 133

```

Db      1818 CATGACCAACCAAAACCAACATCTCTACAGTACACAGCCAGCGTACTCTCGGAGAG 1877
Qy      134 LeuLYaAPaLaGLYVaLaLaThLeuLaLaTPMeVaLaLeMeGLYaSPGLYeunHs 153
Db      1878 CTGAAGATGCGCGCGCTGCGCACCTTGGCTGATGTGATGAATGAAGGATGAGCTGCAC 1937
Qy      154 AsnHeSerAPGlyLeuLaLaLeGLYVaLaLaPheThrgLugLYLeuSeSerGLYLeu 173
Db      1938 AATTTCAGCGATGCGCTGACATGATGCTGCTTTACTGAAGCGTTATCAAGTGTTA 1997
Qy      174 SerThSerVaLaLaValAPheCySHLeGLYLeuPProHLeGLYLeuGLYaSPHeLaLaVal 193
Db      1998 AGTACTCTGTGCTGTGCTGTCTGTGCATAGATGCTGCTCAATGAATTAAGGACTTTCCTT 2057
Qy      194 LeuLeuLYaLaGLYMeThVaLaLYeGLNaLaValLeuLYraAPaLaLeuSeLaLaMet 213
Db      2058 CTACTAAAGCGCTGCATGACCGTTAAGCAGCGCTGCTTATTAATGATTCAGCCATG 2117
Qy      214 LeuLaLaLYLeuGLYMeLaLaThrgLYLaPheHLeGLYHLeGLYLaLaLeuLaValSer 233
Db      2118 CTGCGATCTTGGAAATGACAGAAATTTTCATGCTCATATGCTGAATAATGTTCT 2177
Qy      234 MetTrpLLePheLaLaLeuThraLaGLYLeuPHeMeHLeVaLaLaLeuVaLaSPMeVaLa 253
Db      2178 ATGTGATATTTCACCTTACTGCTGCTTATTCATGATGTTGCTCTGCTGATATGTA 2237
Qy      254 ProGLuMeLeuHLeaAPaLaSerAPHLeGLYCySeSerArGrpGLYTrpPhePhe 273
Db      2238 CCTGAATGCTGCACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2297
Qy      274 LeuGLNaLaLaGLYMeLeuLeuGLYLeuGLYLeuMeLeuLeuLeuLeuLeuLeuLeu 293
Db      2298 TTACAGAAATGCTGGAGATGCTTTGGATTTGAAATGATTAATTCATTAATTTGAA 2357
Qy      294 HisLYsLeuValAPheArGLYLeuAPhe 302
Db      2358 CATAAATCGTCTTCGTATTAATTTTC 2384

```

RESULT 7  
 US-10-769-612-1  
 : Sequence 1, Application US/10769612  
 : Publication No.-US20040141983A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Law, Debbie  
 : APPLICANT: Glsh, Kurt C.  
 : APPLICANT: Murray, Richard  
 : APPLICANT: Culp, Patricia  
 : TITLE OF INVENTION: Compositions Against Cancer Antigen L1V-1 And Uses Thereof  
 : FILE REFERENCE: 05882-0129.CPUS08  
 : CURRENT APPLICATION NUMBER: US/10769, 612  
 : PRIOR FILING DATE: 2004-01-29  
 : PRIOR APPLICATION NUMBER: 60/443, 712  
 : PRIOR FILING DATE: 2003-01-29  
 : PRIOR APPLICATION NUMBER: 09/642, 034  
 : PRIOR FILING DATE: 2000-08-18  
 : PRIOR APPLICATION NUMBER: 09/525, 361  
 : PRIOR FILING DATE: 2000-03-15  
 : PRIOR APPLICATION NUMBER: 09/453, 137  
 : PRIOR FILING DATE: 1999-12-02  
 : PRIOR APPLICATION NUMBER: 09/450, 810  
 : PRIOR FILING DATE: 1999-11-29  
 : PRIOR APPLICATION NUMBER: 09/268, 865  
 : PRIOR FILING DATE: 1999-03-15  
 : NUMBER OF SEQ ID NOS: 18  
 : SOFTWARE: PatentIn version 3.2  
 : SEQ ID NO: 1  
 : LENGTH: 2744  
 : TYPE: DNA  
 : ORGANISM: Homo sapiens  
 : US-10-769-612-1

Alignment Scores:

```

Pred. No.: 1,31e-150 Length: 2744
Score: 1359.50 Matches: 301
Percent Similarity: 40.32% Conservative: 1
Best Local Similarity: 40.19% Mismatches: 0
Query Match: 85.02% Indels: 447
DB: 18 Gaps: 1

US-10-659-004-104 (1-302) x US-10-769-612-1 (1-2744)
Qy      1 MetLaArGLYLeuSeSerVaLaLeuLeuLeuThraPheLaLeuSeSerVaLaThraAPro 20
Db      138 ATGCGAGGAAGTATGCTGATCTTGATCTTGACCTTGACCTTGCCCTGCTGCACAAATCCC 197
Qy      21 LeuHLeGLYLeuLYaLaLaLaPheProGLInThrgLYuLYeLeSePProAsnTrp 40
Db      198 CTTATGAACCTAAAGACGCTGCTTCCCGACCACTGGAATAATTAATGTCGGAATGG 257
Qy      41 GluSeRGLYLeaPNaLaLePLeuLaLaLeSerThraGrGLNLYHLeGLNGLInLeu 60
Db      258 GAATCTGGCATTAATGTTGACTTGGCAATTTCCACGCGCAATATCATCTTCAACAGCTT 317
Qy      61 PheTYraGrGLYGLYLeuAsnAsnSeLeuSeSerVaLaGLYLeuArGLYLeuLeuGLN 80
Db      318 TTCTACCGCTATGAGAAATTAATTTCTTGTCACTGTAAGGGTTCAGAAATTAATCTTCAA 377
Qy      81 AsnLLeGLYLeaAPLeYLeLYaArGLYLeHLeHLeHLeHLeHLeHLeHLeHLeHLeHLeH 100
Db      378 AATTAAGCATATGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 437
Qy      101 AspHLeGLYHLeHLeSeSerAPHLeGLYArGrHLeSeSerAPHLeGLYHLeHLeSeSerAP 119
Db      438 GACCAAGCATCTCAAGACCATGAGCGTCACTCAAGCATGAGCATCACTCAGACCAAC 497
Qy      119 ----- 119
Db      498 GAGCATCACTGACCATTAATCATGCTGCTTGTGTAATAAGCAAAAGCTTTTGC 557
Qy      119 ----- 119
Db      558 CCAGACATGACTCAGATAGTTACGATTAAGATCTTAAGAAACAGCCAGGGAGAGAGCT 617
Qy      119 ----- 119
Db      618 CACCGACCAACAATGCCAGTGTGAAGAAATGCAAGACAGTGTAGTCAAGTAA 677
Qy      119 ----- 119
Db      678 GTGACTCACTGTGTACAACACTGTCTGGAAGAACTCATTTCTAGAGCAATAGAG 737
Qy      119 ----- 119
Db      738 ACTCAAGACTGGAAAACTTCCCAAGATGTAAGCAGCTCCACTCCACCGAGTGC 797
Qy      119 ----- 119
Db      798 ACATCAAGAGCCGGGTGAGCCGCTGCTGTGAGAAAACAATAATCTGTGAGTGA 857
Qy      119 ----- 119
Db      858 CCGGAAAAAGCTTATGATTCAGAAAACAATAATCTCAAGAGTGTTCAT 917
Qy      119 ----- 119
Db      918 GCATCAAGCTACTGATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 977
Qy      119 ----- 119
Db      978 TTCACTATCTGTGTCAGCCATCATCAACCAATGATGATGATGATGATGATGATGATGAT 1037
Qy      119 ----- 119
Db      1038 ACAAGTAAAAAGAGGCTGAATCCCTCAAAAGACTTATTCATAAATAGCTGGGTT 1097

```



Qy 41 GluSerGylleAenValAspLeuAlaIleSerThrArgIlnTyRHisLeuGlnLeu 60  
Db 258 GAATCTGGCATTTAATGTTGACTGGCAATTTCCACGCGCAATTCATCTACCAACAGCTT 317  
Qy 61 PheTyArgTyRgLygluAsnAsnSerLeuSerValgluGlyPheArgLyLeuLeuGln 80  
Db 318 TTCTACCGCTATGAGAAATTAATTTCTTGTCACTGAGAGGCTTCAGAAATTACTTCA 377  
Qy 81 AenIleGlylleAspLyseIleYarGylleHsIleHsIleAspHisAspHisSer 100  
Db 378 AATTTAGGCATAGATTAAGATTAATAAAGATTCATATACCATGACCAAGACATCATCA 437  
Qy 101 AspHisGluHisIleSerAspHisGluArgHisSerAspHisGluHisIleSerAsp 119  
Db 438 GACCAAGCATCACTCAGACCATGAGCGTCATCAGACCATGAGCATCACTCAGACAC 497  
Qy 119 ----- 119  
Db 498 GAGCATCACTGACCATATCATGCTGCTTCTGTAATAAATAGCGAAAGCTTTTGC 557  
Qy 119 ----- 119  
Db 558 CCAAGCATGACTGAGATAGTTCAAGTAAGNTCCTAGAAAAGCCAGAGGAAAGAGACT 617  
Qy 119 ----- 119  
Db 618 CACCGACCAAGACATGCCAGTGTAGAGAAATGTCAGAGACAGTGTAGTGTAGTA 677  
Qy 119 ----- 119  
Db 678 GTGACCTCAACTGTGTACACACTGTCTGTAGAGAACTCACTTTCTAGAGCAATAGAG 737  
Qy 119 ----- 119  
Db 738 ACTCCAAAGACCTGGAAACTCTTCCCAAGATGTAGAGAGTCCACTCCACCCAGTGC 797  
Qy 119 ----- 119  
Db 798 ACATCAAGAGCCGGGTGAGCCGGCTGGCTGTAGAGAAACAAATGAATCTGTGAGTGA 857  
Qy 119 ----- 119  
Db 858 CCCCCAAAAGCTTTATGTATTTCCAGAAACAAATGAATCCTCAGAGTGTTCAT 917  
Qy 119 ----- 119  
Db 918 GCATCAAGACTAGTACATTCATGGGCAATCCAGTTCCGCTGAATGCAACAGAG 977  
Qy 119 ----- 119  
Db 978 TTCAACTATCTCTGTCCAGCATCATCAACAAATGTAGTGTAGATCTTGTCTGATTGAT 1037  
Qy 119 ----- 119  
Db 1038 ACAAGTAAAGAAAGCTGAATCCCTCCAAAGACTATTCATTAATAATAGCTGGGTT 1097  
Qy 119 ----- 119  
Db 1098 GGTGGTTTATAGCATTTGCATCATGATTCCTGTCTGTGCTGGGGTTATCTTAGTG 1157  
Qy 119 ----- 119  
Db 1158 CCTCTCATGAATCGGGTGTCTTTCAAAATTTCTCTGAGTTTCTTTGTGCACTGGCCGTT 1217  
Qy 119 ----- 119  
Db 1218 GGGACTTTGAGTGTGATGCTTTTTCACCTTTCTTCCATTCATGCAAGTACACAC 1277  
Qy 119 ----- 119  
Db 1278 CATAGTCATAGCATGAAGAACAGCAATGGAATGAAGAAAGAGACACTTTTTCAGTCAT 1337

Qy 119 ----- 119  
Db 1338 CTGTCTTCTCAAAACATAGAAAGAAAGTCCATTTTGTATGTCAGCTGGAAGGCTTAACA 1397  
Qy 119 ----- 119  
Db 1398 GCTCTAGAGGCGTGTATTTTCATGTTTCTTGTGAACATGTCCTCATGATGATCAACAA 1457  
Qy 119 ----- 119  
Db 1458 TTTAAGATAGAAAGAAAGAAATCAGAGAAACCTGAAATATGATGATGTGAGATT 1517  
Qy 119 ----- 119  
Db 1518 AAGAAAGATTGTCCAAAGTATGAATTCACACTTTCAACAAATGAGAGAAAGTATATCA 1577  
Qy 119 ----- 119  
Db 1578 GATGATGAACTGAAGCTATTTACAGACACTCACAGAGCCCTCCACTTGTATCT 1637  
Qy 119 ----- 119  
Db 1638 CAGCAGCCTGCAGTCTTGAGAAAGAAAGAGTCAATGATGCTCATGCTACACAGGA 1697  
Qy 119 ----- 119  
Db 1698 GTCTAACAATGAATATGTAACCAAGAGGTGCMAAGATAATGCCATTCACTTCCAGAT 1757  
Qy 119 ----- 119  
Db 1758 ACACCTGGCCAGTCAGAGATCTCATTCACCAACATCATGACTACATATTTCTCAT 1817  
Qy 120 ----- 133  
Db 1818 CATCACCAACCAAAACACATCTCTCACAGTACAGCCAGCCCTACTCTCGGAGAGAG 1877  
Qy 134 LeuYsAspAlaGlyValAlaThrLeuAlaThrMetValIleMetGlyAspGlyLeuHis 153  
Db 1878 CTGAAGATGCGCGCTGCGCATCTGCTGTGATGGATGAATGAGTGGCTGAC 1937  
Qy 154 AsnHisSerAspGlyLeuAlaIleGlyAlaAlaIleAspThrGluGlyLeuSerSerGlyLeu 173  
Db 1938 AATTTACCGCATGCGCTTACGCAATGCTGCTCTTTACTGAAGCTTATCAAGTGTTTA 1997  
Qy 174 SerThrSerValAlaValPheCysHisGluLeuProHisGluLeuGlyAspPheAlaVal 193  
Db 1998 AGTACTCTGTGCTGTGTCTGTCAATGATGCTCATGAATTAGGTGACTTGTCTGTT 2057  
Qy 194 LeuLeuYsAlaGlyMetThrValIleGlnAlaValLeuTyraAsnAlaLeuSerAlaMet 213  
Db 2058 CTACTAAAGCTGGCATGACCGTTTACAGAGGCTGTCTTAAATGCAATGTCCAGCCATG 2117  
Qy 214 LeuAlaTyRleuGlyMetAlaThrGlyIlePheIleGlyHisTyraAlaGluAsnValSer 233  
Db 2118 CTGGCGTATCTTGAATGCAACAGAAATTTTCATTTGCTCTTATGCTGAAAAGTTCCT 2177  
Qy 234 MetTrpIlePheAlaLeuThrAlaGlyLeuPheMetHisValAlaLeuValAspMetVal 253  
Db 2178 ATGTGATATTTGCACTTACTGTGCTGCTTATTCATGATGTGCTGTGATATGCTA 2237  
Qy 254 ProGluMetLeuHisAsnAspAlaSerAspHisGlyCysSerArgTrpGlyTyRPhethe 273  
Db 2238 CCTAAATGCTGCACAAATGATGCTAGTGAACCATGAGATGACCCCTGGGGTATTTCTTT 2297  
Qy 274 LeuGlnAsnAlaGlyMetLeuLeuGlyPheGlyIleMetLeuLeuIleSerIlePheGly 293  
Db 2298 TTACAGAAATGCTGGAGTCTTTTGGGTTTGGATATGTTACTTATTTCCATATTTGAA 2357  
Qy 294 HisAlysIleValPheArgIleAsnPh 302  
Db 2358 CATAAATCGTGTTCGTATTAATTTTC 2384

RESULT 9

```
US-10-198-846-10007
; Sequence 10007, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10007
; LENGTH: 3537
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1, 2, 3533, 3534, 3535, 3536, 3537
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-10007

Alignment Scores:
Pred. No.: 1,928-150 Length: 3537
Score: 1359,50 Matches: 301
Percent Similarity: 40.32% Conservative: 1
Best Local Similarity: 40.19% Mismatches: 0
Query Match: 85.02% Indels: 447
DB: Gaps: 1

US-10-659-004-104 (1-302) x US-10-198-846-10007 (1-3537)

OY 1 MetAlaAaGlyLeuSerValIleLeuIleLeuThrPheAlaLeuSerValThrAspPro 20
DB 138 ATGGGAGAGAGATTTCGTATCTTGATCCGACCTTGGCCCTCTGTGACAAATCC 197
OY 21 LeuHISgluLeuAlaAlaIlePheProGlnThrGluLeuIleSerProAsnTrp 40
DB 198 CTTCACTGAACCTAAAGCAGCTCTTCCCGACACACTGAGAAATTAAGTCCGATTGG 257
OY 41 GluSerGlyIleAsnValAspLeuAlaIleSerThrArgGlnTrpHisLeuGlnLeu 60
DB 258 GAATCTGGCATTAATGTGATGGCAATTTCACACGCGCAATATCATCTCAACAGCTT 317
OY 61 PheTyrArgTyrGlyGluAsnAsnSerLeuSerValIleGluIlePheArgLysLeuGln 80
DB 318 TTCTACCCCTAAGGAAATTAATCTTTGTCAAGTGAAGGTTCAAGAAATTAATCTTCA 377
OY 81 AsnIleGlyIleAspLysIleLysArgIleHisIleHisIleAspHisAspHisSer 100
DB 378 AATTAAGGCATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 437
OY 101 AspHisGluHisHisSerAspHisGluArgHisSerAspHisGluHisHisSerAsp 119
DB 438 GACCGACGAGCATCACTCAGACCATGAGCGTCACTCAGACCATGAGCATCACTCAG 497
OY 119 ----- 119
DB 498 GAGCATCACTTGACCATTAATCATGCTGCTGTGTAATAATTAAGCGAAAGCTTTGC 557
OY 119 ----- 119
DB 558 CCAAGCATGACTCAGATAGTTCAAGTAAAGATCTCTAAGAACAGCCAGGGAAGAGACT 617
OY 119 ----- 119
DB 618 CACCGACCAAGACATGCGATGTGTAGAAAGATGTCAAGACAGTGTAGTGTAGTGA 677
```

```
OY 119 ----- 119
DB 678 GTGACCTCAACTGTGTACAACTGTCTGTGAAGAACTCACTTTAGAGACAATAGAG 737
OY 119 ----- 119
DB 738 ACTCAAGACCTGGAATACTTCCCAAGATGTAGACAGCTCCACTCCACCGAGTGC 797
OY 119 ----- 119
DB 798 ACATCAAGACCGGCGGTGAGCGGCTGCTGTGTAGAAAACAAATGAATCTGTGATGAG 857
OY 119 ----- 119
DB 858 CCGGAAAGGCTTTATGTATTCCAGAAACAAATGAATAATCTCAGAGAGTGTTCAT 917
OY 119 ----- 119
DB 918 GCATCAAGCTACTGACATCTCATGCGATGGCATCCAGGTTCCGCTGAATGCAACAGAG 977
OY 119 ----- 119
DB 978 TTCAACTATCTGTGTCCAGCATCATCAACCAATGTAGTGTATCTGTGTATTCAT 1037
OY 119 ----- 119
DB 1038 ACAAGTGAAGAGGCTGAATTCCTCCAAAGCTATTTCATAAATAGCCTGGGTT 1097
OY 119 ----- 119
DB 1098 GGTGTATTATAGCATTTCATCATCATGATTCCTGTCTGTCTGCGGGTTATCTTAGTG 1157
OY 119 ----- 119
DB 1158 CTTCTCATGAATCGGAGTGTTCCTCAATTTCTCTGAGTTCTGTGTGCACTGGCGGTT 1217
OY 119 ----- 119
DB 1218 GGGACTTGAAGTGATGCTTTTTCACCTTCTTCCACATTTCTCATGCAAGTCAACAC 1277
OY 119 ----- 119
DB 1278 CATAGTCATGACCATGAAGAACCAAGATGAATGAAGAGACCACTTTGATGTCAT 1337
OY 119 ----- 119
DB 1338 CTGTCTTCTCAAAACATAGAAGAAAGTCTATTGATTTCACGTGAAGGCTTAACA 1397
OY 119 ----- 119
DB 1398 GCTTAGAGGCTGTATTTCATGTTTCTTGTGAACATGCTTCACATTGATCAAAACA 1457
OY 119 ----- 119
DB 1458 TTTAAATTAAGAAAGAAAGAAATCAGAGAAACCTGAAATATGATGATGAGATT 1517
OY 119 ----- 119
DB 1518 AAGAGAGATTGCCAAGTATGATTCGAATCTTCAACAAATGAGAGAAAGTAGATACA 1577
OY 119 ----- 119
DB 1578 GATGATGAAGTGAAGCTATTTCAGAGCAGACTCACAAGAGCCCTCCACTTTGATTCT 1637
OY 119 ----- 119
DB 1638 CAGCAGCTGAGTCTTGAAGAAAGAGGTATGATGATGATGATGATGATGATGATGAT 1697
OY 119 ----- 119
DB 1698 GTTCAATGAATATGTATGCCAGAGGCTGCAAGATTAATGATTCATTTCCAGAT 1757
```

```

Oy 119 ----- 119
Db 1758 AACACTGGCGCATGACGATCTCATTCACACCATCATGACTACATCATATTCTCCAT 1817
Oy 120 ----- HiHiPProHISerHisSerGlnArgTyrSerArgGlu 133
Db 1818 CATCACCCACCAACCAACACCATCTCATGATCAGTCCACACCGCTACTCTCGGAGAG 1877
Oy 134 LeuYsAaPaLaGlyValAlaThrLeuAlaTrpMetValIleMetGlyAspGlyLeuHis 153
Db 1878 CTGAAGAGTCCGGGCTGCGCCACTTGCGCTGGATGGATGATGATGGATGGATGGATGGAC 1937
Oy 154 AenPheSerAaPaLaGlyValAlaIleGlyAlaAlaPheThrGluGlyLeuSerSerGlyLeu 173
Db 1938 AATTTCACGATGGCTGATGACATTCGCTGCTTACTGAAAGCTTATCAAGTGTTA 1997
Oy 174 SerThrSerValAlaValAlaPheCysHisGluLeuProHisGluLeuGlyAspPheAlaVal 193
Db 1998 AGTACTCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2057
Oy 194 LeuLeuYsAaPaLaGlyMetThrValIleGlnAlaValLeuYsAaPaLaLeuSerAlaMet 213
Db 2058 CTACTAAAGGCTGGGATGACCGTTAAGCAGCTGCTCTTATTAATGATTCATTCAGCCATG 2117
Oy 214 LeuAlaTyrLeuGlyMetAlaThrGlyIlePheIleGlyHisTyrAlaGlyAsnValSer 233
Db 2118 CTGGCGTATCTTGGAATGCGACAGAGATTTTCATTTGCTCATTTATGCTGAAATGTTTCT 2177
Oy 234 MetTrpIlePheAlaLeuThrAlaGlyLeuPheMetHisValAlaLeuValAspMetVal 253
Db 2178 ATGGGATATTTGCACTTACTGCTGCTTATTCATGATGATGTTGCTGCTGATATGTTA 2237
Oy 254 ProGluMetLeuHisAaPaLaGlyAserAspHisGlyCysSerArgTyrGlyTyrPhePhe 273
Db 2238 CCTAAAGTCTGCAACATGATGCTAGTACCATGATGTAGCCGCTGGGGATTTCTTT 2297
Oy 274 LeuGlnAaPaLaGlyMetLeuGlyPheGlyIleMetLeuLeuIleSerIlePheGlu 293
Db 2298 TTACGAATGCTGGGATCTTTGGGTTTGGAAATTATGTTATTTCCATTTGAA 2357
Oy 294 HisYsIleValPheArgIleAsnPhe 302
Db 2358 CATAAATCGTGTTCGTATAAATTTC 2384

RESULT 10
US-10-188-832-176
; Sequence 176, Application US/10188832
; Publication No: US20040076955A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Aziz, Natabha
; TITLE OF INVENTION: Methods of Diagnosis of Bladder Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Bladder
; FILE REFERENCE: 018501-0023330US
; CURRENT APPLICATION NUMBER: US/10/188, 832
; PRIORITY FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: US 60/302,814
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: US 60/310, 099
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/343, 705
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 176
; LENGTH: 2268
; TYPE: DNA

```

```

/ ORGANISM: Homo sapiens
US-10-188-832-176
Alignment Scores:
Pred. No.: 2,236-150 Length: 2268
Score: 1356.50 Matches: 301
Percent Similarity: 40.00% Conservative: 1
Best Local Similarity: 39.87% Mismatches: 0
Query Match: 84.83% Indels: 453
DB: 17 Gaps: 1

US-10-659-004-104 (1-302) x US-10-188-832-176 (1-2268)
Oy 1 MetAlaArgIleAsnSerValIleLeuIleLeuThrPheAlaLeuSerValThrAsnPro 20
Db 1 ATGGGAGAAAGTATCTGTAATCTTGAATCTGACCTTGGCTCTCTGTCACAAATCC 60
Oy 21 LeuHisGluLeuYsAaPaLaAlaPheProGlnThrThrGlyIleSerProAsnTrp 40
Db 61 CTTCATGACTAAAGCAGCTGCTTCCCGACAGCCACTGAGAAATTAATGTCGAAATTGG 120
Oy 41 GluSerGlyIleAsnValAspPheAlaIleSerThrArgGlnTyrHisLeuGlnGlnLeu 60
Db 121 GAATCTGGCATTAATGTTGACTTGGCAATTCCACAGGCAATATCATCTACACAGCTT 180
Oy 61 PheTyrArgTyrGlyIleAsnAsnSerLeuSerValGluGlyPheArgIleLeuGln 80
Db 181 TTCTACCGCTATGAGAAATTAATTTCTTGTCACTGAAGGGTTCAGAAATTAATCTTCA 240
Oy 81 AsnIleGlyIleAspYsIleYsArgIleHisIleHisHisAspHisAspHisHisSer 100
Db 241 AATATAGGCAATGATTAAGATTAAAGATTCATATACCATGACACGACCACTACCTCA 300
Oy 101 AspHisGluHisHisSerAspHisGluArgHisSerAspHisGluHisHisSerAsp 119
Db 301 GACCACGAGATCACTACGACCATGAGGTCATCAGACCATGAGCATCACTCAGACCA 360
Oy 119 ----- 119
Db 361 GAGCATCACTGACATGATCATCACTCCACCATATATCATGCTTCTGTGTAATAAT 420
Oy 119 ----- 119
Db 421 AAGCGAAAGCTCTTGGCCAGACCATGATCATCAATGTTAGGTAAGATCTTAAGAAC 480
Oy 119 ----- 119
Db 481 AGCAGGGGAAAGAGCTCAACCGACAGAACATGCAAGTGTAGAGAAATGCAAGAC 540
Oy 119 ----- 119
Db 541 AGTGTAGTGTAGTGAAGTGAAGTCAACTGTGTACAAACTGTCTGTGAAGAACTCAC 600
Oy 119 ----- 119
Db 601 TTTTACAGACATAGAGACTCAAGACTGGAAAACTTTCCCAAGATGTAGACAC 660
Oy 119 ----- 119
Db 661 TCACCTCACCCAGTGCATCATCAAAAGCGGGTGAAGCGGCTGGCTGTAGGAAAAACA 720
Oy 119 ----- 119
Db 721 AATGAATGTGAGTGAAGCCCGAAAAAGCTTATGTATTCAGAAACACAAATGAAAT 780
Oy 119 ----- 119
Db 781 CTCAGAGTGTTCATATGATCAAAAGCTACTGACATCTATGAGCATGGGCATCAGGTT 840
Oy 119 ----- 119
Db 841 CCGGTGAATGCAACAGAGTTCACTATCTGTCTCAGGCATCATCAACAAATTGATGCT 900

```

```

QY 119 ----- 119
Db 901 AGATCTGTCTGATTCAATCAAGTGAAGAAAGCTGAATCCCTCCAAAGACCTATTCA 960
QY 119 ----- 119
Db 961 TTACAATAGCGTGGGTTGGTGGTTTAAAGCATTTCATCATCATGTTCTGTCTGTG 1020
QY 119 ----- 119
Db 1021 CTGGGGTTTATTTAGTGCCTCTCATGATCGGGTGTGTTTCAATTTCTCTGAGTTTC 1080
QY 119 ----- 119
Db 1081 CTGTGGCACTGGCCGTTGGGACTTTGAGTGTGATGCTTTTACACCTCTTCCACAT 1140
QY 119 ----- 119
Db 1141 TCTCATGCAAGTCAACCATATGATGATGCAATGAAGAACCAAGCAATGGAATGAAGAA 1200
QY 119 ----- 119
Db 1201 GGACACATTTTGCATCTGTCTCTCAAAAATAGAGAAGAGTGCCTATTGTTGATTC 1260
QY 119 ----- 119
Db 1261 ACGTGAAGGCTTAACAGCTCTAGAGGCTGTATTTTCATGTTTCTTGTGACATGTC 1320
QY 119 ----- 119
Db 1321 CTCACATTTGATCAAAACAATTTAAAGTAAAGAGAAAAAGATCAGAGAAACCTGAAAT 1380
QY 119 ----- 119
Db 1381 GATGATGATGAGATTAAAGAGAGTTTCCAAATGATCTCAATTTCAACAAT 1440
QY 119 ----- 119
Db 1441 GAGGAGAAAGTAAAGATACAGATATGAACTGAAGGCTATTTCGAGCAGACTCACAAG 1500
QY 119 ----- 119
Db 1501 CCGTCCACTTTGATTTCTGAGAGCGCTGAGCTTGAAGAGAAAGAGTATGATGCT 1560
QY 119 ----- 119
Db 1561 CATGCTCATCAAGAGAGTCTACATGATATGTAACCAAGGCTGAAGATTAATGTC 1620
QY 119 ----- 119
Db 1621 CATTCACATTTCCAGATACACTCGGCCAGTCAGAGCATCTCATTCACACCATCATGAC 1680
QY 120 ----- 127
Db 1681 TACCATCATATTCTCATCATCAACACCAAAAACCAATCTTCACAGTCAAGCCGAG 1740
QY 128 ArgTyrSerArgGluGluLeuLeuValAlaValAlaPheCysHisGluLeuProHisGlu 147
Db 1741 CCTTACTCTCGGAGAGAGCTGAAGATGCGCGCTGCGCACCTTTGGCCCTGAGTGGGAA 1800
QY 148 MetGlyAspGlyLeuHisAsnProHisAspGlyLeuAlaAlaGlyValAlaHisThrGlu 167
Db 1801 ATGGGAGTGGCTGACCAATTTCAAGCAATGGCTGACCAATGGTCTCTTTACTGA 1860
QY 168 GlyLeuSerSerGlyLeuSerThrSerValAlaValPheCysHisGluLeuProHisGlu 187
Db 1861 GCGTATCATCAAGGTTTAAAGTCTTGTGCGTGTCTGTGATAGTATGCTCATGA 1920
QY 188 LeuGlyAspPheAlaValLeuLeuValAlaGlyMetThrValIleGlnAlaValLeuTyr 207
Db 1921 TTAGGTGACTTGTCTTCTACTAAAGCTGGCATGACCGTTTAAGCAGGCTGTCTTTAT 1980
QY 208 AsnAlaLeuSerAlaMetLeuAlaTyrLeuGlyMetAlaThrGlyIlePheIleGlyHis 227

```

```

Db 1981 AATGCAATTCAGGCATGCGCGTATCTTGAATGCAACAGGAATTTTCATGTCAT 2040
QY 228 TyrAlaGluAsnValSerMetTrpIlePheAlaLeuThrAlaGlyLeuPheMetHisVal 247
Db 2041 TATGCTAAATGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
QY 248 AlaLeuValAspMetValProGluMetLeuHisAsnAspAlaSerAspHisGlyCysSer 267
Db 2101 GCTTGGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160
QY 268 ArgTrpGlyTyrPhePheLeuGlnAsnAlaGlyMetLeuGlyPheGlyIleMetLeu 287
Db 2161 CGCTGGGGATATTCTTTTAAAGAAATGCTGGATGCTTTGGGTTTGGAAATATATGTTA 2220
QY 288 LeuIleSerIlePheGluHisIleValPheArgIleAsnPro 302
Db 2221 CTATTTTCATATTGGAACATAAATGCTGTGTTGATTAATTTTC 2265

RESULT 11
US-10-659-004-105
; Sequence 105, Application US/10659004
; Publication No. US20050048507A1
; GENERAL INFORMATION:
; APPLICANT: Zhong et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-608
; CURRENT APPLICATION NUMBER: US/10/659, 004
; PRIOR FILING DATE: 2003-09-09
; PRIOR APPLICATION NUMBER: 60/295,607
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/295,661
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/296,404
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/296,418
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/297,414
; PRIOR FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: 60/297,567
; PRIOR FILING DATE: 2001-06-12
; PRIOR APPLICATION NUMBER: 60/298,285
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: 60/298,556
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/299,949
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: 60/300,883
; PRIOR FILING DATE: 2001-06-26
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 187
; SOFTWARE: Curaseqdist version 0.1
; SEQ ID NO 105
; LENGTH: 2268
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2265)
US-10-659-004-105

Alignment Scores:
Pred. No.: 2,236-150 Length: 2268
Score: 1356.50 Matches: 301
Percent Similarity: 40.00% Conservative: 1
Best Local Similarity: 39.87% Mismatches: 0
Query Match: 84.83% Indels: 453
DB: 19 Gaps: 1

US-10-659-004-104 (1-302) x US-10-659-004-105 (1-2268)
1 MetAlaArgLeuSerValIleLeuIleLeuThrPheAlaLeuSerValThrAsnPro 20

```

Db 1 ATGGGAGGAAGTTATCTGTATCTTGATCTTGATCTTGACCTTGGCCCTCTCTGTACAAACCC 60  
QY 21 LeuH1eGluLeuLysAlaAlaAlaPheProGlnThrTyrGluLysIleSerProAsnTrp 40  
Db 61 CTTCAATGAACCTAAAGACAGCTGCTTCCCGACAGACCTGAGAAATTAATGTCGAAATGG 120  
QY 41 GluSerGlyIleAsnValAspLeuAlaIleSerThrArgGlnTyrPheGluGlnLeu 60  
Db 121 GAATCTGGCAATTAAATGTAATGGCAATTTCCACAGCGCAATATCATCTCAACAGAGCTT 180  
QY 61 PheTyrArgTyrGlyLysAsnAsnSerLeuSerValGluGlyPheArgLysLeuGln 80  
Db 181 TTCTACCCCTAATGAGAAATTAATCTTTGTCAAGTTGAGGGGTTGAGAAATTAATTTACTTCA 240  
QY 81 AsnIleGlyIleAspLysIleLysArgIleHisIleHisAspHisAspHisSer 100  
Db 241 AATATAGGCATATGATTAATTAATTAAGAAATCCATATACACACAGACGACCATCATCA 300  
QY 101 AspHisGluHisIleSerAspHisGluArgHisSerAspHisGluHisSerAsp--- 119  
Db 301 GACCAAGACATCTCAGACATGAGCGTCACTGACACATGACATCACTCAGACAC 360  
QY 119 ----- 119  
Db 361 GAGCATCACTGACCATGATCATCATCTCACATTAATCATGCTGCTTCTGTAAAT 420  
QY 119 ----- 119  
Db 421 AAGCGAAAGCTCTTTGCCAGACCATGATCATGATTAAGTTAGTAAGATCTAGAAAC 480  
QY 119 ----- 119  
Db 481 AGCCAGGGGAAAGAGCTCACCGACAGAAATGCCAGTGTAGAAAGAAATGTCAAGAC 540  
QY 119 ----- 119  
Db 541 AGTGTAGTGTAGTGAAGTGAACCTCAACTGTGTATCAACAATGCTCTGAAAGAACTCAC 600  
QY 119 ----- 119  
Db 601 TTTCTAGAGACAAATAGAGATCCAGACCTGGAAACTCTTCCCAAGATGTAGCAGC 660  
QY 119 ----- 119  
Db 661 TCCACTCCACCCAGTGTACATCAAAAGCCGGGTGAGCCGGCTGGTGTAGGAAAA 720  
QY 119 ----- 119  
Db 721 AATGAATCTGTAGTGAAGCCCGAAAGGCTTTATGTATTCAGAAACAACAATGAAT 780  
QY 119 ----- 119  
Db 781 CCTGAGAGTGTTCATATGATCAAAAGCTACTGATCATGGCATGGGATCCAGGTT 840  
QY 119 ----- 119  
Db 841 CCGCTGAATGCAACAGAGTTCAATATCTGTCTCAGCCATCATCAACAAATGATGCT 900  
QY 119 ----- 119  
Db 901 AGATCTGTGATTCATACAAAGTGAAGAAAGGCTGAATTCCTCCAAAGACTATTCA 960  
QY 119 ----- 119  
Db 961 TTACAATAGCCTGGGTGTGTGTTTATAGCCATTTCCATCATCATGATTTCTGTCTGTG 1020  
QY 119 ----- 119  
Db 1021 CTGGGGTTATCTTAGTGCCTCTCATGAATCGGAGTTTTCAAATTTCTCTGAGTTTC 1080  
QY 119 ----- 119  
Db 1081 CTGTGGACACTGGCGGTGGACCTTGAATGTGATGCTTTTTCACACTTCTTCCACAT 1140

QY 119 ----- 119  
Db 1141 TCTCATGCAAGTCAACACATATGATATAGCCATGAAAGAACGACAAATGAAATGAAAA 1200  
QY 119 ----- 119  
Db 1201 GGACCACTTTTATGATCATCTGTCTTCTCAAAACATAGAAAGAAAGTCCATTTTGATTCC 1260  
QY 119 ----- 119  
Db 1261 ACGTGAAGGGTCTTAACAGCTCTAGAGAGCCTGATTTTCATGTTTCTTGTGAACATGTC 1320  
QY 119 ----- 119  
Db 1321 CTCACATTGATCAACAATTTAAAGATAGAAAGAAAGAAATCAAGAGAAACCTGAAAT 1380  
QY 119 ----- 119  
Db 1381 GATGATGATGTGAGATTAAAGAGCATGTTTCCAAATATGATCTCAACTTTCAACAAT 1440  
QY 119 ----- 119  
Db 1441 GAGAGAAAGTATGATACAGATGATCGAACTGAAGGCTATTTACGACAGACTCACAAGAG 1500  
QY 119 ----- 119  
Db 1501 CCTTCCACTTTGATTTCTCAGACAGCTGCACTTTGGAAGAAAGAGTCAATAGACT 1560  
QY 119 ----- 119  
Db 1561 CATGCTATCCACAGAAAGTCTACATGAATATGTACCAGAGGCTGCAAGATTAATGCC 1620  
QY 119 ----- 119  
Db 1621 CATTCACATTTCCAGATACACTGGGCCAGTCAAGCATCTCATTCACACCATCATGAC 1680  
QY 120 ----- 120  
Db 1681 TACCATCATATTTCCATCATCACCAACCAACCACTCCCTCAGATCAGCAGCCAG 1740  
QY 128 ArgTyrSerArgGluGluLeuLysAspAlaGlyValAlaThrLeuAlaTrpMetValIle 147  
Db 1741 CGTACTCTCGGAGAGCTTAAGATGCCGCGCTGCGCACTGTGCTGATGGTGTGTA 1800  
QY 148 MetGlyAspGlyLeuHisAsnPheSerAspGlyLeuAlaIleGlyAlaAlaPheThrGlu 167  
Db 1801 ATGGGTGATGGCTGTGCACAAATTTCAAGCATGGCTTAGCAATTTGGTGTCTTTACTGA 1860  
QY 1861 GGCATTATCAAGTGGTTTAAGTACTTGTGTCTGTGTTCTGTCAATGAGTCCCATGA 1920  
QY 188 LeuGlyAspPheAlaValLeuLeuLysAlaGlyMetThrValLysGlnAlaValLeuLys 207  
Db 1921 TTAGGTGACTTGTGCTGTCTACTAAAGCTGACATGACCGTTAAAGCAGGCTGTCTTAT 1980  
QY 208 AsnAlaLeuSerAlaMetLeuAlaTyrLeuGluLysMetAlaThrGlyIlePheIleGlyHis 227  
Db 1981 AATGCATTGTCAAGCATGCTGCGATATCTTGAATGGCAACGAAATTTTCATGTCAT 2040  
QY 228 TyrAlaGluAsnValSerMetTrpIlePheAlaLeuThrAlaGlyLeuPheMetHisVal 247  
Db 2041 TATGCTGAAAATGTTTCTATGTGATATTTGCATCTTGTGCTGCTTATTCATGATGTT 2100  
QY 248 AlaLeuValAspMetValProGluMetLeuHisAsnAspAlaSerAspHisGlyCysSer 267  
Db 2101 GCTCTGGTGAATAGTACCTGAATGCTGCAACATGATGCTAGTACCATGAGATGTAGC 2160  
QY 268 ArgTyrGlyTyrPhePheLeuGlnAsnAlaGlyMetLeuLeuGlyPheGlyIleMetLeu 287  
Db 2161 CGCTGGGGTATTTCTTTTTCAGAAATGCTGGATGCTTTTGGGTTTGGAAATATATGTTA 2220



```

QY 119 ----- 119
Db 1398 ACGTGAAGGCTTAACAGCTCTAGAGAGCCTGATTTTCATGTTCTTGTGAACATGTC 1457
QY 119 ----- 119
Db 1458 CTCACATTGATCAACAATTTAAGATAGAAGAAAAGATCAGAGAAACCTGAAAT 1517
QY 119 ----- 119
Db 1518 GATGATGATGAGAGATTAGAAGAGAGTTGTCCAAGTATGATCTCACTTTCAACAAT 1577
QY 119 ----- 119
Db 1578 GAGAGAAAGATAGATACAGATGATCGAACTGAAGGCTATTACGAGAGAGCTCAACAGAG 1637
QY 119 ----- 119
Db 1638 CCCTCCCACTTGTGATTTCTAGCAGAGCTGACGTCTTGAAGAAGAGGTATGATAGCT 1697
QY 119 ----- 119
Db 1698 CATGCTCATCCACAGAGAGTCTCAATGATATGTACCCAGAGGGTGCAAGATTAATGC 1757
QY 119 ----- 119
Db 1758 CATTCACTTTCCAGATACACTCGGCGCAGAGCTCAGATCTCATTCAACCACATCATGAC 1817
QY 120 ----- 120
Db 1818 TACCATCATATTTCTCATCATCATCACCACCAACCAACCACTCTCAACAGTCACAGCCAG 1877
QY 128 ArgTyrSerArgGluLeuLysAspAlaGlyValAlaThrLeuAlaTyrMetValIle 147
Db 1878 CGCTACTCTCGGAGAGAGACTGAAGATGCCGCGCTGCCACTTTGGCTGGATGGTATGATA 1937
QY 148 MetGlyAspGlyLeuHisAsnProSerArgPheLeuAlaIleGlyAlaAlaPheThrGlu 167
Db 1938 ATGGGTGATGGCTGCACCAATTTTCAGCGATGCCCTAGCAATTTGGCTGCTTTTACTGAA 1997
QY 168 GlyLeuSerSerGlyLeuSerThrSerValAlaValPheCysHisGluLeuProHisGlu 187
Db 1998 GGCTTATTAAGTGGTTTAAGTACTTCTGTGCTGTGCTGTGCTGATGAGTTGCCCTCATGAA 2057
QY 188 LeuGlyAspPheAlaValLeuLeuLysAlaGlyMetThrValLysGlnAlaValLeuTyr 207
Db 2058 TTAGGTGACTTTGGCTGTTCTAATAAGGCTGGCATGACCGTTAAGCAGGCTGCTTTAT 2117
QY 208 AsnAlaLeuSerAlaMetLeuAlaTyrLeuGlyMetAlaThrGlyIlePheIleGlyHis 227
Db 2118 AATGCAATGTGACGCGCATCTGCGCTATCTTGAATGGCAACAGGAATTTTCATTGGTCAT 2177
QY 228 TyrTlaGluAsnValSerMetTyrIlePheAlaLeuThrAlaGlyLeuPheMetHisVal 247
Db 2178 TATGCTGAATAATGTTTCTATGATGATATTGCACTTACTGCTGCTTATCTCACTGATGAT 2237
QY 248 AlaLeuValAspMetValProGluMetLeuHisAsnAspAlaSerAspHisGlyCysSer 267
Db 2238 GCTCGGTGATATGATGCTGTAATGCTGCAATGATGCTAGTGAACAAGATGATGAC 2297
QY 268 ArgTyrGlyTyrPhePheLeuGlnAsnAlaGlyMetLeuLeuGlyPheGlyIleMetLeu 287
Db 2298 CGCTGGGGGATTTCTTTTTCAGAAATGCTGGAGATGCTTTGGGTTTGGATTAATGTTA 2357
QY 288 LeuIleSerIlePheGluHisIleValPheArgIleAsnPro 302
Db 2358 CTATTTTCATATTTGAACATTAATAATCGTGTGATTAATTTTC 2402

```

RESULT 13  
 US-10-295-027-838  
 ; Sequence 838, Application US/10295027  
 ; Publication No. US2003023250A1

```

GENERAL INFORMATION:
APPLICANT: Afar, Daniel
APPLICANT: Aziz, Natasha
APPLICANT: Ginsberg, Wendy M.
APPLICANT: Glash, Kurt C.
APPLICANT: Glynn, Richard
APPLICANT: Hevezl, Peter A.
APPLICANT: Mack, David H.
APPLICANT: Murray, Richard
APPLICANT: Watson, Susan R.
APPLICANT: Eos Biotechnology, Inc.
TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
FILE REFERENCE: 018501-012500US
CURRENT FILING DATE: US/10/295,027
PRIOR APPLICATION NUMBER: US 09/663,733
PRIOR FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: US 60/350,666
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/335,394
PRIOR FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: US 60/332,464
PRIOR FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: US 60/334,393
PRIOR FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: US 60/340,376
PRIOR FILING DATE: 2001-12-14
PRIOR APPLICATION NUMBER: US 60/347,211
PRIOR FILING DATE: 2002-01-08
PRIOR APPLICATION NUMBER: US 60/347,349
PRIOR FILING DATE: 2002-01-10
PRIOR APPLICATION NUMBER: US 60/355,250
PRIOR FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: US 60/356,714
PRIOR FILING DATE: 2002-02-13
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1386
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 838
LENGTH: 3461
TYPE: DNA
ORGANISM: Homo sapiens
US-10-295-027-838

Alignment Scores:
Pred. No.: 4,24e-150 Length: 3461
Score: 1356.50 Matches: 301
Percent Similarity: 40.00% Conservative: 1
Best Local Similarity: 39.87% Mismatches: 0
Query Match: 84.83% Indels: 453
Gaps: 1

US-10-659-004-104 (1-302) x US-10-295-027-838 (1-3461)
QY 1 MetAlaArgLysLeuSerValIleLeuIleLeuThrPheAlaLeuSerValThrAsnPro 20
Db 138 ATGCGAGGAAGTATCTGATCTGATCTGACCTTGGCTTCTGTCACAAATGCC 197
QY 21 LeuHisGluLeuLysAlaAlaPheProGlnThrGlyLysIleSerProAsnTyr 40
Db 198 CTTATGAACATAAAGCGCTTGTCCCGACCACTGAGAAATTAAGTCCGAATGG 257
QY 41 GluSerGlyIleAsnValAspLeuAlaIleSerThrArgGlnTyrHisLeuGlnGlnLeu 60
Db 258 GAATCGCATTAATGTTGATCTTGCATTTCCACAGGCAATATCATCAACAGCTT 317
QY 61 PheTyrArgTyrGlyGluAsnAsnSerLeuSerValGluGlyPheArgLysLeuGln 80
Db 318 TTCTACCGCTATGAGAAATTAATTTCTTGTCACTTAAGGTTCAAGAAATTACTTCAA 377
QY 81 AsnIleGlyIleAspLysIleLysArgIleHisIleHisAspHisAspHisSer 100

```

Db 378 AATATAGCATAGATTAAGATTTAAAGAAATCCATATACCACTGACCGACCATCATCTCA 437  
 QY 101 Asph1sg1u1sh1sSerAsph1sg1u1rgh1sSerAsph1sg1u1sh1sSerAsp--- 119  
 Db 438 GACCAACGAGATCACTGACCATGAGCGTCACTGACACCATGAGCATCTCAGACCAAC 497  
 QY 119 ----- 119  
 Db 498 GAGCATCTCTGACCATGATCATCTGACCATTAATCATGCTGCTTGTAATAAT 557  
 QY 119 ----- 119  
 Db 558 AAGGAAAAAGCTCTTGCCGACAGACCATGACTCAAGTACTGAGTAAGATCTTAGAAG 617  
 QY 119 ----- 119  
 Db 618 AGCCAGGGGAAAGAGCTCACCGACCAAGAACATGCCAGTGTAGAGAGAAATGTCAGAGAC 677  
 QY 119 ----- 119  
 Db 678 AGTGTAGTGTAGTGAAGTGAACCTCAACTGTGTACAACTGTCTGTGAAGAACTGAC 737  
 QY 119 ----- 119  
 Db 738 TTTCTAGAGCAATAGAGACTCCAAAGACTGGAAAACTCTTCCCAAAAGATGTAGACAGC 797  
 QY 119 ----- 119  
 Db 798 TCCACTCCACCCAGTGTACATCAAAAGACCGGCTGAGCCGGCTGGTGTAGAGAAACA 857  
 QY 119 ----- 119  
 Db 858 AATGAATCTGTAGTGAAGCCCGAAAAAGGCTTTATGTATTCAGAAACAACAATGAAAT 917  
 QY 119 ----- 119  
 Db 918 CCTCAGAGTGTTCATATGATCAAAAGCTACTGACATCTCATGSGATGGGCATCCAGGT 977  
 QY 119 ----- 119  
 Db 978 CCGGTGAATGCAAGAGATTCAACTATCTGTCCAGCATCATCAACCAATGATGCT 1037  
 QY 119 ----- 119  
 Db 1038 AGATCTTGTCTGATTCATCAAGTGAAGAAAGGCTGAATCCCTCCAAAGACTTATCA 1097  
 QY 119 ----- 119  
 Db 1098 TTACAATAGCCTGGGTTGTGTGTTTATAGCCATTTCCATCATGATTCCTGTCTG 1157  
 QY 119 ----- 119  
 Db 1158 CTGGGGGTATCTTAGTGCCCTCATGAAATCGGGTGTGTTTTCAAAATTTCTCTGAGTTTC 1217  
 QY 119 ----- 119  
 Db 1218 CTTTGGAACCTGCGCGTGGACCTTTGAGTGTGATGCTTTTTTACACCTTCTTCACAT 1277  
 QY 119 ----- 119  
 Db 1278 TCTCATGCAAGTACCAACATAGTCTATAGCATGAAAGAACAGCAATGAAATGAAAAA 1337  
 QY 119 ----- 119  
 Db 1338 GGACCACTTTTCACTGATCTGTCTTCTCAAAAATAGAGAAAGTCTATTTTGTATTC 1397  
 QY 119 ----- 119  
 Db 1398 ACGTGAAGGCTTAAACAGCTTAGAGGCGCTGTATTTGATGTTCTTGTGAACATGTC 1457  
 QY 119 ----- 119  
 Db 1458 CTCACATTGATCAACATTTTAAAGTAAAGAGAAAAAGATCAGAGAAACCTGAAAT 1517

QY 119 ----- 119  
 Db 1518 GATGATGATGTGAGATTTAAAGACAGTGTCCAGATATGATCAACTTTTCAACAAT 1577  
 QY 119 ----- 119  
 Db 1578 GAGGAGAAAGTATACATGATGATCGAATGAAAGCTATTTTAGAGCAACTCAGAGAG 1637  
 QY 119 ----- 119  
 Db 1638 CCTCCACTTTGATTTCTCAGACCGCTGAGCTTTGAGAGAGAGAGTCAATGATGCT 1697  
 QY 119 ----- 119  
 Db 1698 CAGTCTATCCACAGAGAGTCTACATGAATATGTATCCACAGAGGCTGCAGAAATTAATGC 1757  
 QY 119 ----- 119  
 Db 1758 CATTCAATTTCCAGATATACATGCGGACGTCAGAGATCTCATTCACCAACCATCATGAC 1817  
 QY 120 ----- 119  
 Db 1818 TACCATCATATTTCTCATCATCAACCAACCAAAACCACTCTCTCACAGTCACAGCCAG 1877  
 QY 128 ArgTyrSerArgGluGluLeuLysAspAlaGlyValAlaPheLeuAlaTrpMetValIle 147  
 Db 1878 CGTACTCTCGGAGAGCTGAAGATGAGCGCGCTGCGCACTTGGCTTGATGTGATGATA 1937  
 QY 148 MetGlyAspGlyLeuHisAsnPheSerAspGlyLeuAlaIleGlyValAlaPheThrGlu 167  
 Db 1938 ATGGGTATGCGCTGCAACATTTTCAAGCATGCGCTTACCAATGTGTGCTTTTACTGAA 1997  
 QY 168 GlyLeuSerSerGlyLeuSerThrSerValAlaValPheCysHisGluLeuProHisGlu 187  
 Db 1998 GGCCTTATCAAGTGGTTTAAGTACTTCTGTGCTGTCTGTCTGTCTGTCTGTCTGTCTGAA 2057  
 QY 188 LeuGlyAspPheAlaValIleLeuLysAlaGlyMetThrValLysGlnAlaValLeuTyr 207  
 Db 2058 TTAGGTGACTTGTGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 2117  
 QY 208 AsnAlaLeuSerAlaMetLeuAlaTyrLeuGlyMetAlaThrGlyIlePheIleGlyHis 227  
 Db 2118 AATGATTTGTACGACCAAGCTGCGCTATCTTGAATGGAACAGGAATTTTCTATGTCAT 2177  
 QY 228 TyrAlaGluAsnValSerMetTrpIlePheAlaLeuThrAlaGlyLeuPheMetHisVal 247  
 Db 2178 TATGCTGAATAATGTTTCTATGTGATATTTGCACTTCTGCTGCTTATTCATGATATGTT 2237  
 QY 248 AlaLeuValAspMetValProGluMetLeuHisAsnAspAlaSerAspHisGlyCysSer 267  
 Db 2238 GCTCGGTGATGATGATGCTGGAATGCTGCAATGATGCAATGATGCTGATGCAATGATGATG 2297  
 QY 268 ArgTrpGlyTyrPhePheLeuGlnAsnAlaGlyMetLeuLeuGlyPheGlyIleMetLeu 287  
 Db 2298 CGCTGGGGGATTTCTTTTACAGATGCTGGAGTGTGTTGGGTTTGGATATATGTTA 2357  
 QY 288 LeuIleSerIlePheGluHisLysIleValPheArgIleAsnPhe 302  
 Db 2358 CTTATTTCCATATTTGAACATTAATCGTGTGTTGATTAATTTTC 2402  
 RESULT 14  
 US-10-295-027-878  
 ; Sequence 878, Application US/10295027  
 ; Publication No. US20030232350A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Afar, Daniel  
 ; APPLICANT: Aziz, Natasha  
 ; APPLICANT: Ginsberg, Wendy M.  
 ; APPLICANT: Gish, Kurt C.  
 ; APPLICANT: Glynn, Richard  
 ; APPLICANT: Hevez, Peter A.  
 ; APPLICANT: Mack, David H.

```
/ APPLICANT: Murray, Richard
/ APPLICANT: Watson, Susan R.
/ TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
/ FILE REFERENCE: 018501-012500US
/ CURRENT APPLICATION NUMBER: US/10/295,027
/ CURRENT FILING DATE: 2002-11-13
/ PRIOR APPLICATION NUMBER: US 09/663,733
/ PRIOR FILING DATE: 2000-09-15
/ PRIOR APPLICATION NUMBER: US 60/350,666
/ PRIOR FILING DATE: 2001-11-13
/ PRIOR APPLICATION NUMBER: US 60/335,394
/ PRIOR FILING DATE: 2001-11-15
/ PRIOR APPLICATION NUMBER: US 60/332,464
/ PRIOR FILING DATE: 2001-11-21
/ PRIOR APPLICATION NUMBER: US 60/334,393
/ PRIOR FILING DATE: 2001-11-29
/ PRIOR APPLICATION NUMBER: US 60/340,376
/ PRIOR FILING DATE: 2001-12-14
/ PRIOR APPLICATION NUMBER: US 60/347,211
/ PRIOR FILING DATE: 2002-01-08
/ PRIOR APPLICATION NUMBER: US 60/347,349
/ PRIOR FILING DATE: 2002-01-10
/ PRIOR APPLICATION NUMBER: US 60/355,250
/ PRIOR FILING DATE: 2002-02-08
/ PRIOR APPLICATION NUMBER: US 60/356,714
/ PRIOR FILING DATE: 2002-02-13
/ Remaining Prior Application data removed - See file wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 1386
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 878
/ LENGTH: 3461
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-295-027-878

Alignment Scores:
Pred. No.: 4,24e-150      Length: 3461
Score: 1356.50           Matches: 301
Percent Similarity: 40.00%      Conservative: 1
Best Local Similarity: 39.87%    Mismatches: 0
Query Match: 84.83%           Indels: 453
DB: 17                   Gaps: 1

US-10-659-004-104 (1-302) x US-10-295-027-878 (1-3461)
Oy 1 MetAaTgLySeLSeSeValIleuTThPheAlaLeuSeValThrAsnPro 20
Db 138 ATGGCGAGAAAGTTATCTGTATCTTGAATCCTGACCTTGGCCCTCTGTGCACAAATCCC 197
Oy 21 LeuHlsgLeuLyAlAlaAheProGlnThrGluLyLeSeSeProAsnTrp 40
Db 198 CTTGATGAACCTAAAGAGCGCTTCCCGGACCACTGAAGAAATTAAGTCCGATTTGG 257
Oy 41 GluSeRgLyIleAsnValAspLeuAlIleSeThraGlnTrpHlLeuGlnGlnLeu 60
Db 258 GAATCTGCATTAATGTTGACTTGCAATTTCCACACGCGCAATATCATCTCAACAGCTT 317
Oy 61 PheTTrArgTyRgLyGluAsnAsnSeLSeSeValGluGlyPheArgLyLeuLengIn 80
Db 318 TTCTAACCGCTATGGAGAAATTAATCTTGTGAGTGAAGGCTTCAGAAATTACTTCAA 377
Oy 81 AsnIleGlyIleAspLyIleLyAArgIleHlslleHlslAspHlAspHlslAspHl 100
Db 378 AATATAGGCATATGATTAATTAAGATTCATATACACATGACACGACCATCATCA 437
Oy 101 AspHlsgLuhHlslSeSeAspHlsgLuhArgHlsgSeAspHlsgLuhHlslSeSeAsp 119
Db 438 GACCAAGCATCACTCAGACCATGAGCGTCACTCAGACCATGAGCATCACTCAGACCAAC 497
Oy 119 ----- 119
```

```
Db 498 GAGCATCACTGTGACATGATCATCACTGCACCATATATATGCTGTTCTGTAAAAAT 557
Oy 119 ----- 119
Db 558 AAGGAAAAGCTCTTGGCCAGACCATGACTCAGATAGTTCCAGTAAAGATCTTGAAGAAC 617
Oy 119 ----- 119
Db 618 AGCCAGGGGAAAAGAGCTCACCCGACGAAACATGCGAGTGTAGAGGAATGTCAAGGAC 677
Oy 119 ----- 119
Db 678 AGTGTATGCTAGTGAAGTGAACCTCAACTGTGTACAAACACTGTCTGTGAAGAACTCAC 737
Oy 119 ----- 119
Db 738 TTCTAGAGCAATAGAGACTCCAAGACTGGAAAACTTTCCCAAGATGTAAAGCAGC 797
Oy 119 ----- 119
Db 798 TCCACTCCACCCAGTGTCACTCAAGAGCCGGGTGAGCGGCTGGCTGTAGGAAAAACA 857
Oy 119 ----- 119
Db 858 AATGAATCTGTGAGTGAAGCCCCGAAAAGCTTTATGTATTCCAGAAACAAATGAAAT 917
Oy 119 ----- 119
Db 918 CCTCAGAGTGTTCATATGATCAAGACTACTGACATCTCATGAGCATGGCATCCAGTT 977
Oy 119 ----- 119
Db 978 CCGCTAATGCAACAGAGTTCAACTATCTGTCCACCATCATCAACAAATGATGCT 1037
Oy 119 ----- 119
Db 1038 AGATCTGTCTGATTCATACAGTGAAAGAAAGGCTGAATCCCTCCAAAGACTATTCA 1097
Oy 119 ----- 119
Db 1098 TTACAAATAGCGTGGTGGTGTATATAGCCATTTCCATCATCACTTCCGTCTGTG 1157
Oy 119 ----- 119
Db 1158 CTGGGGGTATCTTAGTGCCTTCATGATCGGGTGTTCGAATTTCTCTGAGTTTC 1217
Oy 119 ----- 119
Db 1218 CTTGTGGCAGTGGCGGCTTGGGACTTGTAGTGTGATGCTTTTACACCTTCCACAT 1277
Oy 119 ----- 119
Db 1278 TCTCATGCAAGTCAACACCATATGATAGCATGAGAGAACGAGATGAAATGAAAGA 1337
Oy 119 ----- 119
Db 1338 GGAACCTTTTCACTCATCTGTCTTCAAAAACATGAGAAAGTGCTATTTTGATTC 1397
Oy 119 ----- 119
Db 1398 ACGTGAAGGCTTAAACAGCTTAGAGGCGCTGTATTTCATGTTCTTGTGAAACATGTC 1457
Oy 119 ----- 119
Db 1458 CTCACATTGATCAACAATTTAAAGATAGAGAGAAAAAGATCAGAGAACTGAAAT 1517
Oy 119 ----- 119
Db 1518 GATGATGATGTGAGATTAAAGACAGTTGTCCAAAGTATGATCAACTTTCAACAAT 1577
Oy 119 ----- 119
Db 1578 GAGGAGAAAGTATGATACATGATGATCGAACTGAAGGCTTATTAGAGCAGACTCAAGAG 1637
```



```

Db      858 AATGAATCTGTAGTAGAGCCCGAAGAGCTTATGTATTCAGAAACAAATGAAT  917
Qy      119 -----
Db      918 CCTCAGAGTGTTCATATGATCAAGAGTACTGATCATCTGAGCATGCCAGGTT  977
Qy      119 -----
Db      978 CCGCTGAATGCACAGAGTTCAACTATCTGTGTCCAGCCATCATCAACCAATTGATGCT  1037
Qy      119 -----
Db     1038 AGATTTGTCTGTATTCATACAAAGTGAAGAAAGGCTGAATCCCTCCAAAGACTTATTC  1097
Qy      119 -----
Db     1098 TTACAATATGCTGGGTTGTGTGTTTATAGCCATTTCATCATCATGTTTCTGTCTGTG  1157
Qy      119 -----
Db     1158 CTGGGGGTATCTTAGTCCCTCTCATGAATCGGGTGTCTTTCMAATTTCTCTGAGTTTC  1217
Qy      119 -----
Db     1218 CTGTGGCACTGCGCGTTGGGACTTTGAGTGTATGCTTTTTCACCTTCTTCACAT  1277
Qy      119 -----
Db     1278 TCTCATGCAAGTCAACCATATGATGATGATGAAACAGCAATGGAATGAAGA  1337
Qy      119 -----
Db     1338 GGACCACTTTTCAGTCATCTGTCTTCTCAAAACATAGAAAGGCTTATTGATTCC  1397
Qy      119 -----
Db     1398 ACGTGAAGGCTTAACAAGCTCTAGAGGCTGTATTTCATGTCTTGTGTAACATGTC  1457
Qy      119 -----
Db     1458 CTCACATTGATCAACAATTTAAAGATAGAAAGAAAGAAATCAGAGAAACCTGAAT  1517
Qy      119 -----
Db     1518 GATGATGATGTGAGATTAGAAAGCATGTGTCAAATGATCTCAACTTTCACAAAT  1577
Qy      119 -----
Db     1578 GAGGAGAAATGATACAGATGATGAACTGAAGGCTATTTCAGACAGACTCACAGAG  1637
Qy      119 -----
Db     1638 CCCTCCCACTTTGATTCTCAGCAGCCTGACATCTTGAAGAAGAGGTCAATGACT  1697
Qy      119 -----
Db     1698 CATGCTCATCCACAGAAAGTCTCAATGATATGTACCAGAGGTGCAAGATTAATGC  1757
Qy      119 -----
Db     1758 CATTCACTTTCCAGATACACTCGGCCAGTCAGACGATTCATTCACCAACCATCATGAC  1817
Qy      120 -----
Db     1818 TACCATCATATTTCTCATCATCACACCAACCAAAACCAACATCTCACAGTCCAGCCAG  1877
Qy      128 ARGTYTSEPAAGGIGIULeuLYaSPALaGIYValAlaThrLeuAlaTrpMetValIle  147
Db     1878 CGCTACTCTCGGAGAGAGCTGAAGATGCCGCCACTTTCGCTGAGTGTGATA  1937
Qy      148 MetGlyAaSPGlyLeuHIaSPaSPheSerAaSPGlyLeuAlaIleGlyAlaAlaPheThrGlu  167

```

```

Db      1938 ATGGGTATGGCTGCAACATTTTCAGAGATGGCCCTAGCAATTTGTCCTTTTACTGAA  1997
Qy      168 GlyLeuSerSerGlyLeuSerThrSerValAlaValPheCysHisLeuLeuProHisGlu  187
Db      1998 GGTATATCAAGTGGTTTAAAGTACTTCTGTTCTGTGTCTGTCAATGAGCTTCATGAA  2057
Qy      188 LeuGlyAaSPheAlaValLeuLeuLYaSPALaGIYMetThrValLYSGlnAlaValLeuTyr  207
Db      2058 TTAGTGACTTTGCTGTCTACTAAAGCTGACATGACCGATTAGCAGGCTGTCTTAT  2117
Qy      208 AsnAlaLeuSerAlaMetLeuAlaTyrLeuGlyMetAlaThrGlyIlePheIleGlyHis  227
Db      2118 AATGATTTGTACGCATGCTGGCGTATCTTGGAAATGCAACAGAAATTTTCATTGTCTAT  2177
Qy      228 TyrAlaGluAaSPaValSerMetTrpIlePheAlaLeuThrAlaGlyLeuPheMetHisVal  247
Db      2178 TATGCTGAATAAGTTTCTATGTGATATTCACATTACTGCTGCTGTATTCATGTATGTT  2237
Qy      248 AlaLeuValAaSPMetValProGluMetLeuHisAaSPAlaSerAaSPHisGlyCysSer  267
Db      2238 GCTCTGTTGATATGTACTGAAATGCTGCACAAATGATGCTAGTACCATGATGATGAC  2297
Qy      268 ArgTrpGlyTyrPhePheLeuGlnAsnAlaGlyMetLeuGlyPheGlyIleMetLeu  287
Db      2298 CGCTGGGGTATTTCTTTTACAGAAATGCTGGATGCTTTGGGTTTGGAAATTAATGTTA  2357
Qy      288 LeuIleSerIlePheGluHisIleValIlePheArgIleAaSPhe  302
Db      2358 CTATATTTCATATTTGAACATAAATCGTGTTCGTATTAATTTTC  2402

```

Search completed: March 20, 2005, 01:51:06  
 Job time : 677 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OW protein - nucleic search, using frame\_plus\_p2n model

Run on: March 19, 2005, 22:52:01 ; Search time 213 Seconds

(without alignment)  
2319.979 Million cell updates/sec

Title: US-10-659-004-104

Perfect score: 1599

Sequence: 1 MARKSVLTLTFLSVTNP.....FGIMLLISFEKIVRINF 302

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 1202784 seqs, 81813359 residues

Total number of hits satisfying chosen parameters: 240568

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=xlh  
-O=/cgn2\_1/USPRO.spool/US10659004/runat\_14032005\_181236.15771/app.query.fasta\_1.455  
-DB=Issued Patents NA -QEMT=fastcap -SUFFIX=rn1 -MINMATCH=0.1 -LOOFC=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTPRM=prco -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10659004@cgn2\_1.1.69.@runat\_14032005\_181236.15771 -NCEU=6 -ICPU=3  
-NO MMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV TIMEOUT=130 -WARN TIMEOUT=10 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Issued Patents NA:

1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/Backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1356.5	84.8	2268	US-09-642-034-4	Sequence 4, Appli
2	1356.5	84.8	3461	US-09-642-034-1	Sequence 1, Appli
3	959	60.0	1310	US-08-311-023-1	Sequence 1, Appli
4	959	60.0	2404	US-08-311-023-3	Sequence 3, Appli
5	439	27.5	4573	US-09-814-915A-29	Sequence 29, Appli
6	395	24.7	2229	US-09-774-528-283	Sequence 283, App
7	374	23.4	1888	US-09-599-360B-38	Sequence 38, Appli
8	283	17.7	464	US-09-621-976-1342	Sequence 1342, Ap
9	212.5	13.3	633	US-09-270-767-2100	Sequence 2100, Ap
10	212.5	13.3	633	US-09-270-767-17382	Sequence 17382, A
11	209.5	13.1	754	US-09-270-767-1621	Sequence 1621, Ap
12	209.5	13.1	754	US-09-270-767-16903	Sequence 16903, A

13	161	10.1	522	US-09-663-600A-149	Sequence 149, App
14	159	9.9	894	US-09-489-039A-5631	Sequence 5631, Ap
15	156	9.8	1203	US-09-543-681A-1979	Sequence 1979, Ap
16	152	9.5	584	US-09-663-600A-55	Sequence 55, Appli
17	143	8.9	10304	US-09-627-465B-1	Sequence 1, Appli
18	143	8.9	34230	US-09-949-016-12052	Sequence 12052, A
19	143	8.9	128470	US-09-949-016-13765	Sequence 13765, A
20	142	8.9	377	US-08-332-766A-1	Sequence 1, Appli
21	140.5	8.8	825	US-09-711-164-198	Sequence 198, App
22	140.5	8.8	825	US-09-711-164-236	Sequence 236, App
23	140.5	8.8	1340	US-09-513-999C-14928	Sequence 14928, A
24	140.5	8.8	1771	US-09-949-016-4307	Sequence 4307, Ap
25	137	8.6	789	US-09-902-540-8863	Sequence 8863, Ap
26	137	8.6	8540	US-09-902-540-944	Sequence 944, App
27	137	8.6	767677	US-09-949-016-12147	Sequence 12147, A
28	137	8.6	767677	US-09-949-016-17361	Sequence 17361, A
29	136	8.5	315	US-09-949-016-6385	Sequence 6385, Ap
30	135.5	8.5	97195	US-09-949-016-12212	Sequence 12212, A
31	135.5	8.5	97196	US-09-949-016-16971	Sequence 16971, A
32	133.5	8.3	765	US-09-902-540-8591	Sequence 8591, Ap
33	133.5	8.3	6682	US-09-902-540-906	Sequence 906, App
34	133.5	8.3	35154	US-09-949-016-12384	Sequence 12384, A
35	133.5	8.3	39154	US-09-949-016-12801	Sequence 12801, A
36	133.5	8.3	39443	US-09-949-016-14326	Sequence 14326, A
37	133.5	8.3	39443	US-09-949-016-14327	Sequence 14327, A
38	133.5	8.3	64610	US-09-949-016-12214	Sequence 12214, A
39	132.5	8.3	1769	US-09-428-636-3	Sequence 3, Appli
40	132	8.3	30656	US-09-949-016-14613	Sequence 14613, A
41	131	8.2	492	US-09-902-540-6150	Sequence 6150, Ap
42	131	8.2	1479	US-09-902-540-297	Sequence 297, App
43	131	8.2	27933	US-09-949-016-12369	Sequence 12369, A
44	131	8.2	142783	US-09-949-016-15127	Sequence 15127, A
45	131	8.2	767677	US-09-949-016-12147	Sequence 12147, A

#### ALIGNMENTS

RESULT 1  
US-09-642-034-4  
Sequence 4, Application US/09642034  
Patent No. 6762020  
GENERAL INFORMATION:  
APPLICANT: Mack, David  
APPLICANT: Gish, Kurt C.  
APPLICANT: Eos Biotechnology, Inc.  
TITLE OF INVENTION: No. 6762020e1 Methods of Diagnosing Breast Cancer,  
TITLE OF INVENTION: Compositions, and Methods of Screening for Breast  
FILE REFERENCE: 018501-009700US  
CURRENT APPLICATION NUMBER: US/09/642,034  
CURRENT FILING DATE: 2000-08-18  
PRIOR APPLICATION NUMBER: US 09/268,865  
PRIOR FILING DATE: 1999-03-15  
PRIOR APPLICATION NUMBER: US 09/450,810  
PRIOR FILING DATE: 1999-11-29  
PRIOR APPLICATION NUMBER: US 09/453,137  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: US 09/525,361  
PRIOR FILING DATE: 2000-03-15  
PRIOR APPLICATION NUMBER: WO PCT/US00/06952  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 4  
LENGTH: 2268  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: open reading frame encoding human breast cancer  
US-09-642-034-4  
Alignment Scores:

Pred. No.: 4.79e-148 Length: 2268  
 Score: 1356.50 Matches: 301  
 Percent Similarity: 40.00% Conservative: 1  
 Best Local Similarity: 39.87% Mismatches: 0  
 Query Match: 84.93% Indels: 453  
 DB: 4 Gaps: 1

US-10-659-004-104 (1-302) x US-09-642-034-4 (1-2268)

```

Oy 1 MetAlaArgLysLeuSerValIleuLleuThrPheAlaLeuSerValThrAsnPro 20
Db 1 ATGGGAGAGAGATTATCTGTAATCTTGATCCGACTTGGCCCTCTCTGACAAATCCC 60
Oy 21 LeuHisGluLeuValAlaAlaPheProGlnThrGluValIleSerProAsnTrp 40
Db 61 CTTCATGAACATAAAGCAGCTGCTTCCCGACACACTGAGAAATTAAGTCCGATTTGG 120
Oy 41 GluSerGlyIleAsnValAspLeuAlaIleSerThrArgGlnTyrHisLeuGlnLeu 60
Db 121 GAATCTGGCATTTAATGTGAAGCTTGGCAATTTCCACAGCGCAATATCATCTACACAGCTT 180
Oy 61 PheTyrArgTyrGlyGluAsnAsnSerLeuSerValGluGlyPheArgLysLeuLeuGln 80
Db 181 TTCTACCCGCTATGGAGAAAATAATTTCTTGTCAGTTGAAGGCTCAGAAAATTAATTCTCA 240
Oy 81 AsnIleGlyIleAspLysIleLysArgIleHisIleHisIleAspHisAspHisIleSer 100
Db 241 AATATAGGCATATGATAGATTAAAGAAATCCATATACACCATGACACAGACCATCATCA 300
Oy 101 AspHisGluHisIleSerAspHisIleGluArgHisSerAspHisIleHisSerAsp--- 119
Db 301 GACCAGAGCATCACTCAGACCATAGAGCTCACTCAGACCATGACATCATCATCAGACCA 360
Oy 119 ----- 119
Db 361 GAGCATCACTGACCATGATCATCATCTCAACATAATCATGCTGTTCTGTAAAAAT 420
Oy 119 ----- 119
Db 421 AAGCAGAAAGCTCTTTGCCAGACCATGATCATGATGATTCAAGTAAAGATCTAGAAAC 480
Oy 119 ----- 119
Db 481 AGCCAGGGGAAAGAGCTCACCAGACCAAGCAATGCCAGTGTAGAAAGAAATCTAAGGAC 540
Oy 119 ----- 119
Db 541 AGTGTAGTCTAGTGAAGTGAAGCTCAACTGTGTAAACAACATGTCTGTGAAGAACTGAC 600
Oy 119 ----- 119
Db 601 TTTCTAGAGACAATAGAGACTCCAAAGACCTGGAAAACTTTCCCAAGATGTAAAGCAGC 660
Oy 119 ----- 119
Db 661 TTCACCTCCACCCAGTGTCACTCAAAAGAGCCGGGTGAGCCGGCTGGCTGTAGAAAA 720
Oy 119 ----- 119
Db 721 AATGATCTGTAGTGAAGCCCGAAAAAGCTTTATGTATTTCCAGAAACAAATGAAAT 780
Oy 119 ----- 119
Db 781 CCTCAGAGTGTTCATATGATCAAAAGTACTGACATCTCATGGCATGGGATCCAGGTT 840
Oy 119 ----- 119
Db 841 CCGCTGAATGCAACAGAGTTCATCTCTGTCCAGCCATCATCAACCAATGATGCT 900
Oy 119 ----- 119
Db 901 AGATCTGTCTGATTCATACAAAGTAAAGAAAGGCTGAATCCCTCCAAAGACCTATTCA 960
  
```

```

Oy 119 ----- 119
Db 961 TTACAAATAGCCGTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1020
Oy 119 ----- 119
Db 1021 CTGGGGGTATCTTAGTGCCTCTCATGAATCGGCTGTTTTCAAATTTCTCTGAGTTTC 1080
Oy 119 ----- 119
Db 119 ----- 119
Oy 1081 CTGTGGCACTGGCCGTTGGACTTTGAGTGTGATGCTTTTTCACCTTTCCACAT 1140
Oy 119 ----- 119
Db 1141 TCTCATGCAAGTCAACACCATATGATATAGCATGAGAACAGCAATGGAATGAAAA 1200
Oy 119 ----- 119
Db 1201 GGACCACTTTTCACTCATCTGTCTTTCAAAACATAGAAAGAGCCTATTGATTCC 1260
Oy 119 ----- 119
Db 1261 ACGTGAAGGCTCTAGACGCTCTAGAGGCGCTGATTCATGTTCTTTGAAACATGTC 1320
Oy 119 ----- 119
Db 1321 CTCACATTGATCAACAATTTAAAGATTAAGAAAGAAAGAAATCAGAAAGAACTGAAAT 1380
Oy 119 ----- 119
Db 1381 GATGATGATGTGAGATTAAAGAACAGTGTCCAAGTATGATCTCACTTTCAACAAT 1440
Oy 119 ----- 119
Db 1441 GAGGAGAAAGTAGATACAGATGATGAACTGAGAGCTATTATTCAGACAGACTCACAGAG 1500
Oy 119 ----- 119
Db 1501 CCTTCCACTTGTGATTCAGACAGCTGACAGTCTTTGAAAGAAAGAGTCAATGATGCT 1560
Oy 119 ----- 119
Db 1561 CATGCTCATCAAGAAAGTCTACAAATGATATGATCCAGAGGCTGCAAGAAATATGC 1620
Oy 119 ----- 119
Db 1621 CATTCACATTTCCAGATACAGTCCGCGCATGACAGATCTCATTCACCAACATATGAC 1680
Oy 120 ----- 120
Db 1681 TACCATATTTCTTCATCATCATCACCAACCAACCAACATCTTCACAGTACAGCCAG 1740
Oy 128 ArgTyrSerArgGluGluLeuLysAspAlaGlyValAlaThrLeuAlaTrpMetValIle 147
Db 1741 CCGTACTCTCGGAGAGAGCTGAAGATCCGCGCTGCGCACTTTGGCTGTGATGTGATA 1800
Oy 148 MetGlyAspArgLysLeuHisAsnPheSerAspArgLysLeuAlaIleGlyValAlaPheThrGlu 167
Db 1801 ATGGGTGAATGGCTGGCAAAATTTCCAGCGATGCTTGAAGATTTGGCTGCTTTTCTGAA 1860
Oy 168 GlyLeuSerSerGlyLeuSerThrSerValAlaValPheCysHisGlyLeuProHisGlu 187
Db 1861 GGCTTATCAAGTGTTAAGTACTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1920
Oy 188 LeuGlyAspPheAlaValIleuLeuValIleGlyMetThrValIleGlnAlaValLeuTyr 207
Db 1921 TTAGGTGACTTGTGCTGTTCTACTAAAGCTGGCATGAGCCGTTAAACAGGCTGCTTTAT 1980
Oy 208 AsnAlaLeuSerAlaMetLeuAlaTyrLeuGlyMetAlaThrGlyIlePheIleGlyHis 227
Db 1981 AATGCAATGTGACGCGCATGCTGGCGTATCTTGAAGGCAACGGAATTTTCAATTTGCTCAT 2040
Oy 228 TyrAlaGluAsnValSerMetTrpIlePheAlaLeuThrAlaGlyLeuPheMetHisVal 247
  
```

```

Db      2041  TATGCTGAAAATGTTCTATGATGGATATTTGCACTTACGCTGGCTTATTCATGTATGTT 2100
Oy      248  AIAeuVAJaBPMeVaiProGIuWecLeuHisaAeAPalaSeAAPHiAGLYCySseR 267
Db      2101  GCTCGGTGATATGTAATGTAACCTGAATATGCTGCACAAATGATGCTAGAGACCAATGATGAC 2166
Oy      268  AATGTGTGYTYrPhePheLeuGlnAaNaAgiWecLeuLeuGlyPheGlyTlWecLeu 287
Db      2161  CGCTGGGGGATTTCTTTTTCAGAAATCTGGAGCTGTTGGGCTTTGGAAATATGTTA 2220
Oy      288  LeuISeSerIlePheGlnHileuYelValPheArgGileanPhe 302
Db      2221  CTATATTCATATTTGAACATPAAATCGTGTTCATPAAATTTTC 2265

RESULT 2
US-09-642-034-1
; Sequence 1, Application US/09642034
; Patent No. 6762020
; GENERAL INFORMATION:
; APPLICANT: Mack, David
; APPLICANT: Gish, Kurt C.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: No. 6762020el Methods of Diagnosing Breast Cancer,
; TITLE OF INVENTION: Compositions, and Methods of Screening for Breast
; TITLE OF INVENTION: Cancer Modulators
; FILE REFERENCE: 018501-009700US
; CURRENT APPLICATION NUMBER: US/09/642,034
; CURRENT FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: US 09/268,865
; PRIOR FILING DATE: 1999-03-15
; PRIOR APPLICATION NUMBER: US 09/450,810
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: US 09/453,137
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: US 09/525,361
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: WO PCT/US00/06952
; PRIOR FILING DATE: 2000-03-15
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 3461
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human breast cancer protein BCR4 CDNA
; NAME/KEY: CDS
; LOCATION: (138)..(2405)
; OTHER INFORMATION: human BCR4
US-09-642-034-1

Alignment Scores:
Pred. No.: 9,246-148 Length: 3461
Score: 1356.50 Matches: 301
Best Similarity: 40.00% Conservative: 1
Best Local Similarity: 39.87% Mismatches: 0
Query Match: 84.83% Indels: 453
DB: 4 Gaps: 1

US-10-659-004-104 (1-302) x US-09-642-034-1 (1-3461)
Oy      1  MetAlaAAGLYleuSeValIleLeuIleuThrPheAlaLeuSeValThrAnPro 20
Db      138  ATGGGAGAAAGTTATCTGTATCTGTGAATCTGACCTTGGCCCTCTGTGCACAAATCC 197
Oy      21  LeuHISGluLeuYAlaAlaIaPheProGlnThrGluYIleSeProAnTrp 40
Db      198  CTTGATGAACATAAAAGAGCTGCTTTCCGCCAGACCACTGAGAAATTAAGTCCGAATTGG 257
Oy      41  GUGSeGlyIleAenValaPheAlaIleSeThrArgGlnTyHISleuGlnLeu 60

```

Db	258	GAATTCGGCATTAATGTTGACTTGGCAATTTCCACAGCGCAATATCATCTCAACACAGTT	317
Qy	61	PheTyrArgTyrGlyGluAsnAsnSerLeuSerValGluGlyPheArgIysLeuLeuGln	80
Db	318	TTCTACCGCTATGGAGAAATATATCTTGTTCAGTTGAAGGTTTCAGAAATTTACTTCA	377
Qy	81	AsnIleGlyIleAspIysIleIysArgIleIleIleIleHisIleHisAspHisAspHisHisSer	100
Db	378	AATATAGCGCATAGATTAAGATTMAAAGATTCATATACCCATGACACAGACCATCATCTCA	437
Qy	101	AspHisGluHisHisSerAspHisGluIysHisSerAspHisGluHisHisSerAsp---	119
Db	438	GAACACGACATCACTCTGACCACTAGAGCGGTCACTAGACCATGAGCATCATCTACAGCAC	497
Qy	119	-----	119
Db	498	GAGCATCATCTGACCATGATCATCATCTCACCATATCATGTCGTTGGTAAAT	557
Qy	119	-----	119
Db	558	AAGCGAAAAGCTCTTGGCCAGACCATGACTCAGTATGTTCAAGTAAAGATCTAGAAAC	617
Qy	119	-----	119
Db	618	AGCCAGGGGAAAAGAGCTCACCGACAGACATGCGCACTGCTAGAGAAATGTCAAGAC	677
Qy	119	-----	119
Db	678	AGTGTAGTGTAGTAAAGTAACTGCTCAACTGTGTACAACAAGTGTCTGAAGAACTCAC	737
Qy	119	-----	119
Db	738	TTTCTAGAGACAATAGAGACTCCAAAGCTGGAAAACCTCTCCCAAGATGTAAAGCAC	797
Qy	119	-----	119
Db	798	TCCACTCCACCAGTGTCACTCAAAGACCGGGGTAGCGCGTGGTGTAGGAAAAACA	857
Qy	119	-----	119
Db	858	AATGAATCTGATGATGAGACCCCGAAAAGCTTTATGTATTCACAAAACAAATGAAT	917
Qy	119	-----	119
Db	918	CCTCAGAGTGTTCATATGATCAAAAGCTATGACATCTCATGCGATGGCATCCAGGTT	977
Qy	119	-----	119
Db	978	CCGCTGAATGCACAGAGTTCATCTCTGCCAGCATCATCAACCAATTGATGCT	1037
Qy	119	-----	119
Db	1038	AGATCTGTCTGATTCATACAAAGTAAAGAAAGGCTGAATCCCTCCAAAGACTATTCA	1097
Qy	119	-----	119
Db	1098	TTACAAATAGCTCGGGTGTGTGTTTATAGCAATTTCCATCATCAAGTTCTCTGTCTG	1157
Qy	119	-----	119
Db	1158	CTGGGGGTTATCTTAGTGCCTTCATGATCGGGGTGTTTCAAAATTTCTCTGAGTTTC	1217
Qy	119	-----	119
Db	1218	CTTGAGCACTGGCCGTTGGGACTTTGAGTGTATGCTTTTACACCTTCTTCACAT	1277
Qy	119	-----	119
Db	1278	TCTCATGAGATGCACACCATAGTCATAGCATGAAGAACAGCAATGAATGAATAAGA	1337
Qy	119	-----	119
Db	1338	GGACCACTTTTCAGTCATCTGTCTTCTCAAAACATAGAAAGAGTGCCTATTTTGATTC	1397

QY	119	-----	119
Db	1398	ACGTGAAGGCTCTTAACAGCTTAGAGAGCCTGATTTTCATGTTTCTTGTGAAACATGTC	1457
QY	119	-----	119
Db	1458	CTCAATTGATCAAAACAATTAAAGATAGAAAGAAAGAAATCAGAGAAACCTGAAT	1517
QY	119	-----	119
Db	1518	GATGATGATGAGATTAAAGACAGTGTCTCCAGTATGAATCTCACTTTCAACAAAT	1577
QY	119	-----	119
Db	1578	GAGAGAAAGTAGATACAGATGATCGAAGCTAAAGGCTATTTACAGACAGACTCAAGAG	1637
QY	119	-----	119
Db	1638	CCCTCCCACTTGATTTCTCAGACAGCCTGACGCTTGGAAGAGAGAGGTATGATAGCT	1697
QY	119	-----	119
Db	1698	CATGCTCATCCACAGAGTCTACATGAATATGTACCCAGAGGTGCAAGATTAATGC	1757
QY	119	-----	119
Db	1758	CATTACATTTCCACAGATACACTGGCGACGTACAGACATCTCATTCACACCATCATGAC	1817
QY	120	-----HbHbProHbSserHbSergin	127
Db	1818	TACCATCATATTTCTCATCATCATCACACCAACCAACCATCTCTACAGTACAGCCG	1877
QY	128	ATGTYTserATgGluGluLeuLeuYAspAlaGlyValAlaThrLeuAlaTrpMetValIle	147
Db	1878	CCCTACTCTCCGGAAGAGACTAAAGATGCGGCGCTGCCACTTTGGCCTGATGATGATA	1937
QY	148	MetGlyAspGlyLeuHbAAsnPheserAspGlyLeuAlaIleGlyAlaAlaaphethrglu	167
Db	1938	ATGGGTGATGGCTCTCACAAATTTACAGAGATGGCTTACCAATTTGGTGTCTTTACTGAA	1997
QY	168	GlyLeuSerSerglyLeuSerThrservalAlaValPheCyHbSgIUleuProHbSgu	187
Db	1998	GGCTTATCAAGTGGTTTAAAGTACTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGAA	2057
QY	188	LeuGlyAspPhaAlaValLeuLeuYAlaGlyMetThrValIysGlnAlaValLeuTyx	207
Db	2058	TTAGGTGACTTTGCTGTTCTACTTAAAGCTGGCATACCGTTAAGCAGGCTGCTTTAT	2117
QY	208	AsnAlaLeuSerAlaMetLeuAlaTyLeuGlyMetAlaThrGlyIlePheIleGlyHbS	227
Db	2118	AATGCAATTTCCAGCCATGCTGCGATCTTGTGAAATGGCAACAGAAATTTTCATTGGTCAAT	2177
QY	228	TyrAlaGluAsnValSerMetTrpIlePheAlaLeuThrAlaGlyLeuPheMetHbSval	247
Db	2178	TATGCTGAAGAAATGTTTCTATGTGATATTTGGACTTACTGTGCTTATTCATGATAGTAT	2237
QY	248	AlaLeuValAspMetValProGluMetLeuHbIsaAsnAspAlaSerAspHbSgIleCySer	267
Db	2238	GCTCTGGTTGATATGTATGTAACCTGAATGCTGCACAAATGATGTAAGCATGATGATAGC	2297
QY	268	ArgTirpGlyTyxPhePheLeuGlnAsnAlaGlyMetLeuLeuGlyPheGlyIleMetLeu	287
Db	2298	CGCTGGGGGGAATTTCTTTTACAGATGCTGGAGATGCTTTGGGTTTGGATATATGTTA	2357
QY	288	LeuIleSerIlePheGluHbSlyArgIleValPheArgIleAsnPro	302
Db	2358	CTTATTTCCATATTTGAACATAAATCGTGTTCGTATTAATTTTC	2402

```

GENERAL INFORMATION:
APPLICANT: MANNING, David Lockwood
APPLICANT: NICHOLSON, Robert Ian
APPLICANT: GEE, Julia Margaret
APPLICANT: GREEN, Christopher Douglas
TITLE OF INVENTION: METHODS FOR PREDICTING THE BEHAVIOUR OF
TITLE OF INVENTION: BREAST TUMOURS
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Young & Thompson
STREET: 745 South 21st Street
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/311,023
FILING DATE: 22-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: PATCH, Andrew J.
REGISTRATION NUMBER: Reg. No. 5693465 32,925
REFERENCE/DOCKET NUMBER: WCM.56
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703/521-2297
TELEFAX: 703/685-0573
TELEX: 248425

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1310 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1293

US-08-311-023-1

Alignment Scores:
Pred. No.: 5.32e-102 Length: 1310
Score: 959.00 Matches: 194
Percent Similarity: 90.41% Conservative: 4
Best Local Similarity: 88.58% Mismatches: 19
Query Match: 59.97% Indels: 5
DB: 1 Gaps: 1

US-10-659-004-104 (1-302) x US-08-311-023-1 (1-1310)

QY 86 Lvsllelylsvatlllehlshlhshlsarphlsarphshlsesrphsh-----glu 1033
   ::::| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 629 ACAAATAATGCCATTCACATTTC-CACGATACACTCGCCAGTCAACAGCATTCATTAC 687
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 104 Hshhsesrphshlsagluarghshsesrphshlglnshhsesrphshshprshs 1233
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 688 CACCATCATGACTCATCATCATATTTCTCCATCATCACACACCAAAACACCATCTGCAC 747
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 124 Serhisserglnargtyrserrrgglugluleulyasrpallagluyalatrleula 1433
Db 748 AGTCACAGCAGGCGTACTCTCGGAGGAGCTGAAGAAGTGCCGGCGTCCGACATTTGGCC 807
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 144 Trpmecvalillemetglyarpglyleuhlsanphsesrpsglyleualilegilyala 1633
Db 808 TGGAATGTGATTAATGGGTATGGCTGCACAATTTCAAGCAATGGCTTAGCAATTGGTCT 867
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 164 AlaphethglugluleuserserqylueSerThSerValAlavAlpheCyvhiSglu 1833
Db 868 GCATTTTCAAGGCTTATCAAGTGGTTTAAGACTTCTGTGCTTCTGATGAG 927
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Oy	184	LeuProHISGluLeuGIYAspPheAlaValLeuLeuValAGlyMetThrValIleGln	203
Db	928	TTGGCTCTAGTAATTAAGTGACTTTGGCTGTCTTCTACTAAAGGCTGGCATACCGTTAAGAG	987
Oy	204	AlaValLeuTYrAsnAlaLeuSerAlaMetLeuAlaTYrLeuGIYMetAlaThrGIYIle	223
Db	988	GCTGTCCTTTAATATGATTCATGTGCAGCCATGCTGGCGTATCTTGGAAATGGCAACGAAAT	1044
Oy	224	PheIleGIYHISTYrAlaGluAsnValSerMetTPIIlePheAlaLeuThrAlaGIYLeu	243
Db	1048	TTCAATGGCTCATTAATGCTGAATAATGTTCTATGTGATTAATTTGACCTTACGCTGGCTTA	1107
Oy	244	PheMetHISValAlaLeuValAspMetValProGIYMetLeuHISAsnAspAlaSerAsp	263
Db	1108	TTTCATGTATGTGCTGCTGGTGTGATTTGGTACTGAAGTGCACAAAGATGCTGTATGAC	1167
Oy	264	HISGIYCyseAspArgTIRGILTYrPhePheLeuGlnAsnAlaGIYMetLeuLeuGIYIle	283
Db	1168	CATGATGATTAACCCGCTGGGGGATTTCTTTTACGAATGCTGGGATCTCTTTGGGTTTT	1227
Oy	284	GIYIleMetLeuLeuIleSerIlePheGluHISLysIleValAlaPheArgIleAsnPhe	302
Db	1228	GGAATTAGTACTTAT-TTCCATATTGAGCATTAATACGAGTTCGT-ATTAATTTTC	1282

```

1      RESULT 4
2      US-08-311-023-3
3      Sequence 3, Application US/08311023
4      Patent No. 5693465
5      GENERAL INFORMATION:
6      APPLICANT: MANNING, David Lockwood
7      APPLICANT: NICHOLSON, Robert Ian
8      APPLICANT: GEE, Julia Margaret
9      APPLICANT: GREEN, Christopher Douglas
10     TITLE OF INVENTION: METHODS FOR PREDICTING THE BEHAVIOR
11     TITLE OF INVENTION: BREAST TUMOURS
12     NUMBER OF SEQUENCES: 3
13     CORRESPONDENCE ADDRESS:
14     ADDRESSEE: Young & Thompson
15     STREET: 745 South 23rd Street
16     CITY: Arlington
17     STATE: VA
18     COUNTRY: USA
19     ZIP: 22202
20     COMPUTER READABLE FORM:
21     MEDIUM TYPE: Floppy disk
22     COMPUTER: IBM PC compatible
23     OPERATING SYSTEM: PC-DOS/MS-DOS
24     SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
25     CURRENT APPLICATION DATA:
26     APPLICATION NUMBER: US/08/311,023
27     FILING DATE: 22-SEP-1994
28     ATTORNEY/AGENT INFORMATION:
29     NAME: PATCH, Andrew J.
30     REGISTRATION NUMBER: Reg. No. 5693465 32,925
31     REFERENCE/DOCKET NUMBER: WCM.56
32     TELECOMMUNICATION INFORMATION:
33     TELEPHONE: 703/521-2297
34     TELEFAX: 703/685-0573
35     TELEX: 248425
36     INFORMATION FOR SEQ ID NO: 3:
37     SEQUENCE CHARACTERISTICS:
38     LENGTH: 2404 base pairs
39     TYPE: nucleic acid
40     STRANDEDNESS: double
41     TOPOLOGY: linear
42     MOLECULE TYPE: CDNA
43     US-08-311-023-3

```

Alignment Scores:	
Pred. No.:	1,37e-101
Score:	959.00
Percent Similarity:	90.41%
Best Local Similarity:	88.58%
Length:	2404
Matches:	194
Conservative:	4
Mismatches:	19

Query Match:	59.97%	Indels:	5
DB:	1	Gaps:	1
US-10-659-004-104 (1-302) x US-08-311-023-3 (1-240)			

OY		86	LysIleValArgLLeHisIleHisIleHisAspHisAspHisIleSerAspHisIle-----Glu	103
Dd		629	AGAAATMAATGCCATTACACTTTC-CACGATACACTCGGGCCAGTAGCATCTCACTTACC	687
OY		104	HisHisSerAspHisGluAArgHisSerAspHisGluHisHisSerAspHisHisProHis	123
Dd		688	CACCATCAGATGATCATCATATTTCTTCATCATCACCAACCAAAACACCATCTTGAC	747
OY		124	SerHisSerGlaArgGyrSerArgGluLeuLeuAspAlaGluValAlaThrLeuAla	143
Dd		748	AGTCACAGGCACCGCTACTCTCGGAGAGAGACTGAAGAATGCGCGCTGCCACTTTGGCC	807
OY		144	TrpMetValIleMetGlyAspGlyLeuHisAsnPheSerAspGlyLeuAlaIleGlyAla	163
Dd		808	TGGATGTGATATAATGGGTATGGCTGCACAAATTCAGCGATGGCCGTACGAATGGTGT	867
OY		164	AlaPheThrGluGlyLeuSerSerGlyLeuSerTherValAlaAlaPheCysHisGlu	183
Dd		868	GCTTTTACTGAAGGCTTAATCAAGTGGTTTAAGTACTCTGTGCTGTGTTCTGCATGAG	927
OY		184	LeuProHisGluLeuGlyAspPheAlaValLeuLeuLysAlaGlyMetThrValLysGln	203
Dd		928	TTGCCCTCATGAATTAAGTGAATCTTGGCTGTTCTACTAAAGCTGGCAGATGACCGTTAAGCAG	987
OY		204	AlaValLeuTyrrAsnAlaLeuSerAlaMetLeuAlaTyrrLeuGlyMetAlaThrGlyIle	223
Dd		988	GCTGTCCTTTATGATCATGTCAGCCCATGCTGGCGATGCTTGGAAATGGCAACAGAAATT	1047
OY		224	PheIleGlyHisTyrrAlaGluAsnValSerMetTroIlePheAlaLeuThrAlaGlyLeu	243
Dd		1048	TTTCATTGGTCATTATGCTGAAAATGTTTTCTATGGATGAAATTTGGACTTACGTGGCTTAA	1107
OY		244	PheMetHisValAlaLeuValAspMetValProGluMetLeuHisAsnAspAlaSerAsp	263
Dd		1108	TTCAATGTATGTTGCTCTGGTTGATATGGTACCTGAATGCGCAACAAGATGCTATGAC	1167
OY		264	HisGlyCysSerArgTrpGlyTyrrPhePheLeuGlnAsnAlaGlyMetLeuLeuGlyPhe	283
Dd		1168	CATGATGTAGCGCGCTGGGGGTATTTCTTTTTCAGAGATGCTGGGATGCTTTTGGGTTTT	1227
OY		284	GlyIleMetLeuLeuLeuSerIlePheGluHisGlyValLevalPheArgIleAspPhe	302
Dd		1228	GGAAATTAATGTAACCTAAT-TCCATATTTGAACATMAAAATCGGTGCT-ATAAAATTC	1282

```

RESULT 5
US-09-814-915A-29
; Sequence 29, Application US/09814915A
; Patent No. 6750015
; GENERAL INFORMATION:
; APPLICANT: Horwitz, Kathryn
; APPLICANT: Richer, Jennifer
; TITLE OF INVENTION: Progesterone Receptor-Regulated Gene Expression and Methods Relat
; TITLE OF INVENTION: Thereeto
; FILE REFERENCE: 2848-39
; CURRENT APPLICATION NUMBER: US/09/814,915A
; CURRENT FILING DATE: 2002-03-21
; PRIOR APPLICATION NUMBER: 60/214,870
; PRIOR FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29
; LENGTH: 4573
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-814-915A-29

```

Alignment Scores: 1.96e-40 Length: 4573  
Pred. No.:



APPLICANT: Dumas Milne Edwards, J.B.  
APPLICANT: Bouguetere, L.  
APPLICANT: Jobert, S.  
TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides  
FILE REFERENCE: GENSET.050CP3  
CURRENT APPLICATION NUMBER: US/09/599,160B  
PRIOR FILING DATE: 2000-06-21  
PRIOR APPLICATION NUMBER: 60/113,686  
PRIOR FILING DATE: 1998-12-22  
PRIOR APPLICATION NUMBER: 60/141,032  
PRIOR FILING DATE: 1999-06-25  
PRIOR APPLICATION NUMBER: 09/469,099  
PRIOR FILING DATE: 1999-12-21  
NUMBER OF SEQ ID NOS: 123  
SOFTWARE: Patent.pm  
SEQ ID NO 38  
LENGTH: 1888  
TYPE: DNA  
ORGANISM: Homo Sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 139..1389  
NAME/KEY: sig\_peptide  
LOCATION: 139..1198  
OTHER INFORMATION: Von Heijne matrix  
OTHER INFORMATION: score 5.00  
OTHER INFORMATION: seq HLLAGFCWVVLG/VV  
NAME/KEY: polyA\_signal  
LOCATION: 1854..1859  
NAME/KEY: polyA\_site  
LOCATION: 1873..1888  
US-09-599-360B-38

Alignment Scores:  
Pred. No.: 1,92e-33 Length: 1888  
Score: 374.00 Matches: 81  
Percent Similarity: 62.05% Conservative: 40  
Best Local Similarity: 41.54% Mismatches: 52  
Query Match: 23.39% Indels: 23  
Gaps: 4

US-10-659-004-104 (1-302) x US-09-599-360B-38 (1-1888)

QY 109 GUAAGHISERAPRHIGLU-----HSHISERAPRHIEPHROHISERHIS 125  
DB 1204 GAGAAAGACACCCAGCACCCAGCTGTGCCCTCTGGCCACCAAGCCACAGCTAT 1263  
QY 126 SerGlnArgTyrSerArgGluGluLeuValAspAlaValAlaThrLeuAlaTrpMet 145  
DB 1264 GGGCACCCG-----GGTGGCATGATATACCTGGATG 1296  
QY 146 ValIleMetGlyAspGlyLeuHisAsnAspSerAspGlyLeuAlaIleGlyAlaAlaPhe 165  
DB 1297 GTCCCTCGGAGATAGTCTACACAACTCACTAGTGGCTGGCCATGAGTGTCTTC 1356  
QY 166 ThrGluGlyLeuSerSerGlyLeuSerThrSerValAlaValPheCysHisGluLeuPro 185  
DB 1357 TCTATGCTTCTCTCCG-CGGCTCAGTACCACTTACGGGCTTCTGTCATGAGCTGCC 1415  
QY 186 HisGluLeuGlyAspPheAlaValIleuLeuValAspGlyMetThrValIleGluAlaVal 205  
DB 1416 CACCAACTGGGTGACTTGTGCTGCTGCTCCAGTACAGGCTGTCTTTTGGGGGCTGTG 1475  
QY 206 LeuTyrAsnAlaLeuSerAlaMetLeuAlaTyrLeuGlyMetAlaThrGlyIlePheIle 225  
DB 1476 CTGCTGACCTCGTCTCGAGCCCTGGGATGGGGGTGCACTCTGGGGGTGGGGCTC 1535  
QY 226 GlyHisTyrAlaGluAsnValSerMetTrpIlePheAlaLeuThrAlaGlyLeuPheMet 245  
DB 1536 AGCCTGGCCCTGTCCCTCACTCCCTGGGTGTGGGTGACAGCTGCGGCTCTTCTC 1595  
QY 246 HisValAlaLeuValAspMetValProGluMetLeuHisAsnAsp-----AlaSer 262  
DB 261 HisValAlaLeuValAspMetValProGluMetLeuHisAsnAsp-----AlaSer 262

DB 1596 TATGAGCCCTTGTGACATGCTACCAAGCCCTGTTGCTCCGAGCCCTGCTACG 1655  
QY 263 AspHisGlyCysSerArgTrpGlyTyrPhePheLeuGlnAsnAlaGlyMetLeuLeuGly 282  
DB 1656 CCCCAT-----GTGCTCTGCGAGGGGCTGGGGCTGTCTGGGG 1694  
QY 283 PheGlyIleMetLeuLeuIleSerIlePheGluHisIleVal 297  
DB 1695 GCGGCTCATGCTTGCATACCTGCTGGAGGAGGCTACTG 1739

RESULT 8  
US-09-621-976-1342  
Sequence 1342, Application US/09621976  
Patent No. 6639063  
GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J.B.  
APPLICANT: Jobert, S.  
APPLICANT: Giordano, J.Y.  
TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
FILE REFERENCE: GENSET.054PR2  
CURRENT APPLICATION NUMBER: US/09/621,976  
CURRENT FILING DATE: 2000-07-21  
NUMBER OF SEQ ID NOS: 19335  
SOFTWARE: Patent.pm  
SEQ ID NO 1342  
LENGTH: 464  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 92..412  
NAME/KEY: sig\_peptide  
LOCATION: 92..160  
OTHER INFORMATION: Von Heijne matrix  
OTHER INFORMATION: score 12  
OTHER INFORMATION: seq RLLLSLVSGALG/LG  
NAME/KEY: misc\_feature  
LOCATION: 325  
OTHER INFORMATION: n=a, g, c or t

Alignment Scores:  
Pred. No.: 9.17e-24 Length: 464  
Score: 283.00 Matches: 58  
Percent Similarity: 65.94% Conservative: 33  
Best Local Similarity: 42.03% Mismatches: 37  
Query Match: 17.70% Indels: 10  
Gaps: 2

US-10-659-004-104 (1-302) x US-09-621-976-1342 (1-464)

QY 163 AlaAlaPheThrGluGlyLeuSerSerGlyLeuSerThrSerValAlaValPheCysHis 182  
DB 2 GCTGCTTCTGTGATGCTTCTCAGGGGCTCAGTACCACTTACAGGCTTCTGCAAT 61  
QY 183 GluLeuProHisGluLeuGlyAspPheAlaValIleuLeuValAspGlyMetThrValys 202  
DB 62 GAGCTGCCCACTAAGTGGGTGACTTGCATGCTGCTCCAGCTCAGGCTGCTTGG 121  
QY 203 GlnAlaValLeuTyrAsnAlaLeuSerAlaMetLeuAlaTyrLeuGlyMetAlaThrGly 222  
DB 122 CGGCTGCTGCTGAGAGCTGTGTCTGAGCCCTGGGATGGGGGTGAGCTCTGGGG 181  
QY 223 IlePheIleGlyHisTyrAlaGluAsnValSerMetTrpIlePheAlaLeuThrAlaGly 242  
DB 182 GTGGGGCTCAGCCGAGGCGCTGTCCCTCACTCCCTGGGTGTGGGTGACAGCTGCTGG 241  
QY 243 LeuPheMetHisValAlaLeuValAspMetValProGluMetLeuHisAsnAsp----- 260  
DB 242 GTCTTCTCTATGTGGGCTTGTGACATGCTACCAAGCCCTGCTTCTGCTCCGAGGCC 301  
QY 261 ---AlaSerAspHisGlyCysSerArgTrpGlyTyrPhePheLeuGlnAsnAlaGlyMet 279  
DB 261 ---AlaSerAspHisGlyCysSerArgTrpGlyTyrPhePheLeuGlnAsnAlaGlyMet 279

```

Db      302 CTGCTTACGCCCCAT-----GTGCTCTTCACAGGCGCTGGGCTG 340
QY      280 LeuGluGlyPheGlyIleMetLeuLeuIleSerIlePheGluHisIleVal 297
      341 CTGCTGGGGGGGGCGCTCATCTTGCCATAACCTCGCTGAGAGCGGCTACTG 394

RESULT 9
US-09-270-767-2100
; Sequence 2100, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2100
; LENGTH: 633
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-2100

Alignment Scores:
Pred. No.:      2,546-15      Length:      633
Score:          212.50      Matches:      47
Percent Similarity: 65.22%      Conservative: 28
Best Local Similarity: 40.87%      Mismatches: 29
Query Match:    13.29%      Indels:      11
DB:             4           Gaps:          4

US-10-659-004-104 (1-302) x US-09-270-767-2100 (1-633)
QY      187 GluLeuGlyAAPPheAlaValLeuLeuYsaIleGlyMetThrValIleGlnAlaValLeu 206
      3 GAACGTGGGGGACTTGTCTCTGTGCTCCAAACAGGTGTCTCATCGGAAGCCCTCAT 62
QY      207 TyrAsnAlaLeuSerAlaMetLeuLeuIleGlyMetAlaThrGlyIlePheIleGly 226
      63 ATGAACATTGTAGCTCACTGCTTGAAGCTTTGGGCGCATGCGGGGCTGTATTGCT 122
QY      227 HisTyrAlaGluAsnValSerMetTrrIlePheAlaLeuThrAlaGlyLeuPheMetHis 246
      123 GGAATTGGAGACGGCATGCCAATGATTTATGACGCCGCTGGTCTTCTCTGAT 182
QY      247 ValAlaLeuValAspMetValProGluMet-----LeuHisAsnAsp-----AlaSer 262
      183 ATTGCTTTTGGCCGATCTGTGCTCCACAAATGATGTGGCCCAATCCGAAATGGCCAAA 242
QY      263 Asp---HisGlyCyseSerArgTrrpGlyTyrPhePheLeuGlnAsnAlaGlyMetLeuLeu 281
      243 GATCCAAAGGCT-----ATATAATATACAAATTTTGTGATCTTCTG 284
QY      282 GlyPheGlyIleMetLeuLeuIleSerIlePheGluHisIleVal 296
      285 GGTCGACTGATATATGTGGCTATGTGCCCTCATGAACAGCATTTA 329
DB

RESULT 10
US-09-270-767-17382
; Sequence 17382, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 17382
; LENGTH: 633
; TYPE: DNA

```

```

; ORGANISM: Drosophila melanogaster
US-09-270-767-17382

Alignment Scores:
Pred. No.:      2,546-15      Length:      633
Score:          212.50      Matches:      47
Percent Similarity: 65.22%      Conservative: 28
Best Local Similarity: 40.87%      Mismatches: 29
Query Match:    13.29%      Indels:      11
DB:             4           Gaps:          4

US-10-659-004-104 (1-302) x US-09-270-767-17382 (1-633)
QY      187 GluLeuGlyAAPPheAlaValLeuLeuYsaIleGlyMetThrValIleGlnAlaValLeu 206
      3 GAACGTGGGGGACTTGTCTCTGTGCTCCAAACAGGTGTCTCATCGGAAGCCCTCAT 62
QY      207 TyrAsnAlaLeuSerAlaMetLeuLeuIleGlyMetAlaThrGlyIlePheIleGly 226
      63 ATGAACATTGTAGCTCACTGCTTGAAGCTTTGGGCGCATGCGGGGCTGTATTGCT 122
QY      227 HisTyrAlaGluAsnValSerMetTrrIlePheAlaLeuThrAlaGlyLeuPheMetHis 246
      123 GGAATTGGAGACGGCATGCCAATGATTTATGACGCCGCTGGTCTTCTCTGAT 182
QY      247 ValAlaLeuValAspMetValProGluMet-----LeuHisAsnAsp-----AlaSer 262
      183 ATTGCTTTTGGCCGATCTGTGCTCCACAAATGATGTGGCCCAATCCGAAATGGCCAAA 242
QY      263 Asp---HisGlyCyseSerArgTrrpGlyTyrPhePheLeuGlnAsnAlaGlyMetLeuLeu 281
      243 GATCCAAAGGCT-----ATATAATATACAAATTTTGTGATCTTCTG 284
QY      282 GlyPheGlyIleMetLeuLeuIleSerIlePheGluHisIleVal 296
      285 GGTCGACTGATATATGTGGCTATGTGCCCTCATGAACAGCATTTA 329
DB

RESULT 11
US-09-270-767-1621
; Sequence 1621, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1621
; LENGTH: 754
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-1621

Alignment Scores:
Pred. No.:      7,466-15      Length:      754
Score:          209.50      Matches:      46
Percent Similarity: 51.22%      Conservative: 38
Best Local Similarity: 28.05%      Mismatches: 69
Query Match:    13.10%      Indels:      11
DB:             4           Gaps:          3

US-10-659-004-104 (1-302) x US-09-270-767-1621 (1-754)
QY      135 LysAspAlaGlyValAlaThrLeuAlaTrrpMetValIleMetGlyAspGlyLeuHisAsn 154
      1 AAGGAACAGCGGACGAGAGGTGGCTGTATGAACCTTGGCCAACTCAATGACAT 60
QY      155 PheSerAspGlyLeuAlaIleGlyValAlaPheThrGluGlyLeuSerSerIleLeuSer 174
      61 TTCACACAGCGTCTAGCCGTGGCTGATCTTTGTGTCTTCAAGACAGCGCATCTTA 120

```

```
Qy 175 ThseValAlaValAPheCyNHleuProHleuLeuGlyAAspPheAlaValLeu 194
    ::::::::::::::::::::
Db 121 GCTACTTTGGCAATATGCTTCAATGAAATTCGACGAGGAGGATTTGGCAATCTCG 180
Qy 195 LeuValAlaGlyMetThr-----ValYsglnAlaValLeuTyAAsnAlaLeuSer 211
    ::::::::::::::::::::
Db 181 CTTGCATCCGGATTCAGTCGCTGGACCGCCGCGCTGGCAATCTACACGCGCGGAGCT 240
Qy 212 AlameleuAlaTyLeuGlyMetAlaThrGlyLeuPheIleGlyHleTyAlaGlyAsn 231
    ::::::::::::::::::::
Db 241 GCGCTGCTGGCTCTTATGATGACATCGGAGCTCCGCGTAACTCGCGCATGAGGCA 300
Qy 232 ValserMetTrIlePheAlaLeuThrAlaGlyLeuPheMetHleValAlaLeuValAsp 251
    ::::::::::::::::::::
Db 301 CGTACTTCGTGGATTTATGCGCTTCAACGCGCGCTTCTGCAATTCCTGCTGTCACA 360
Qy 252 MetValProGlyMetLeuHleAAsnAlaSerAspHleTyGlySerArgTrpGlyTyx 271
    ::::::::::::::::::::
Db 361 GTATTACCTGATCTCTTGAAGAGAGAGGACGCAAGATCCATTAAG----- 408
Qy 272 PhePheLeuGlnAAsnAlaGlyMetLeuLeuGlyPheGlyIleMetLeuIleSer--- 290
    ::::::::::::::::::::
Db 409 -----CAGCTGCTAGCACTGGTATTGTCATTCGCTTAATGCGCGTATGACCATG 459
Qy 291 IlePheGlyHle 294
    ::::::::::::::::::::
Db 460 CTATTCCAAACAC 471

RESULT 12
US-09-270-767-16903
; Sequence 16903, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homurger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1998-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 16903
; LENGTH: 754
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-16903

Alignment Scores:
Pred. No.: 7,46e-15 Length: 754
Score: 209.50 Matches: 46
Percent Similarity: 51.22% Conservative: 38
Best Local Similarity: 28.05% Mismatches: 69
Query Match: 13.10% Indels: 11
DB: 4 Gaps: 3

US-10-659-004-104 (1-302) x US-09-270-767-16903 (1-754)
Qy 135 LyseAraAlaGlyValAlaThrLeuAlaTrpMetValIleMetGlyAAspGlyLeuHleAsn 154
    ::::::::::::::::::::
Db 1 AAGGAAACAGCCGACGAAAGTGGCTGGTATCTGAACCTTGGCCAACTCAATTGACAAAT 60
Qy 155 PheSerAspGlyLeuAlaIleGlyAlaAlaPheThrGlyLeuSerSerGlyLeuSer 174
    ::::::::::::::::::::
Db 61 TTCACACACGGGTCTAGCGGTGGCTGATCTTTTGGTGTCTTCACACACGGGATCTTA 120
Qy 175 ThseValAlaValAPheCyNHleuProHleuLeuGlyAAspPheAlaValLeu 194
    ::::::::::::::::::::
Db 121 GCTACTTTGGCAATATGCTTCAATGAAATTCGACGAGGAGGATTTGGCAATCTCG 180
Qy 195 LeuValAlaGlyMetThr-----ValYsglnAlaValLeuTyAAsnAlaLeuSer 211
    ::::::::::::::::::::
Db 181 CTTGCATCCGGATTCAGTCGCTGGACCGCCGCGCTGGCAATCTACACGCGCGGAGCT 240
Qy 212 AlameleuAlaTyLeuGlyMetAlaThrGlyLeuPheIleGlyHleTyAlaGlyAsn 231
```

```
Db 241 GCGCTGCTGGCTCTTATGATGACATCGGAGGCTCCGCGGCTAAAGCTCGGCATGAGGCA 300
Qy 232 ValserMetTrIlePheAlaLeuThrAlaGlyLeuPheMetHleValAlaLeuValAsp 251
    ::::::::::::::::::::
Db 301 CGTACTTCGTGGATTTATGCGCTTCAACGCGCGCTTCTGCAATTCCTGCTGTCACA 360
Qy 252 MetValProGlyMetLeuHleAAsnAlaSerAspHleTyGlySerArgTrpGlyTyx 271
    ::::::::::::::::::::
Db 361 GTATTACCTGATCTCTTGAAGAGAGAGGACGCAAGATCCATTAAG----- 408
Qy 272 PhePheLeuGlnAAsnAlaGlyMetLeuLeuGlyPheGlyIleMetLeuIleSer--- 290
    ::::::::::::::::::::
Db 409 -----CAGCTGCTAGCACTGGTATTGTCATTCGCTTAATGCGCGTATGACCATG 459
Qy 291 IlePheGlyHle 294
    ::::::::::::::::::::
Db 460 CTATTCCAAACAC 471

RESULT 13
US-09-663-600A-149
; Sequence 149, Application US/09663600A
; Patent No. 6573068
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Duclert, Aymeric
; TITLE OF INVENTION: EXTENDED CDNAS FOR SECRETED PROTEINS
; FILE REFERENCE: 31.US3.CIP
; CURRENT APPLICATION NUMBER: US/09/663,600A
; CURRENT FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/191,997
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: 60/066,677
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/069,957
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/074,121
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/081,563
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/096,116
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: 60/099,273
; PRIOR FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 229
; SOFTWARE: Patent .pm
; SEQ ID NO 149
; LENGTH: 522
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: polyA site
; LOCATION: 512..522
; NAME/KEY: misc_feature
; LOCATION: 11
; OTHER INFORMATION: n=a, g, c or t
US-09-663-600A-149

Alignment Scores:
Pred. No.: 1.94e-09 Length: 522
Score: 161.00 Matches: 47
Percent Similarity: 50.36% Conservative: 23
Best Local Similarity: 33.81% Mismatches: 62
Query Match: 10.07% Indels: 7
DB: 4 Gaps: 3

US-10-659-004-104 (1-302) x US-09-663-600A-149 (1-522)
Qy 1 MetAlaArglyLeuSerValIleLeuIleuThrPheAlaLeuSerValThrAsnPro 20
    ::::::::::::::::::::
Db 84 ATGCACACAAATAATTTGCTTCATTTGTTGCTGACATTTATTTTTCATTCATTCACATCAT 143
```





**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - nucleic search, using frame\_p1uc\_model

Run on: March 19, 2005, 21:58:34 ; Search time 4918 Seconds  
(without alignments)  
2975.493 Million cell updates/sec

Title: US-10-659-004-104  
1599

Perfect score: 1599  
Sequence: 1 MARLSTVILITFALSVTNP.....FGIMLISIFBKIVRINF 302

## Scoring table:

BLOSUM62  
Xgapop 10.0 , Ygapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

## Command line parameters:

-MODEL=frame+g2n.model -DEV=x1h  
-Q=/cgm2\_1/USPTO.spool/US10659004/runat\_14032005\_181235\_15741/app\_query.fasta\_1.455  
-DB=GenEmbl -CPMT=faaescap -SUFFIX=rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=b1h -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USBR=US10659004.OCGN\_1\_1\_3731 @runat\_14032005\_181235\_15741 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEJUDERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

## Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_hcg:\*  
3: gb\_in:\*  
4: gb\_on:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_scs:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1359.5	85.0	2744	6	CQ834422 Sequence
2	1359.5	85.0	2744	6	CQ854099 Sequence
3	1359.5	85.0	2744	9	HSU41060 Sequence
4	1359.5	85.0	3523	6	CQ493869 Sequence

5	1356.5	84.8	3461	6	AX465588 Sequence
6	1356.5	84.8	3461	6	AX829136 Sequence
7	1352.5	84.6	3445	6	AX511618 Sequence
8	1349.5	84.4	2776	6	AX207207 Sequence
9	1324.5	82.8	2265	6	CQ854111 Sequence
10	1311.5	82.0	3461	6	AX207205 Sequence
11	1237.5	77.4	3610	10	BC055012 Sequence
12	1148	71.8	3287	10	AB071697 Sequence
13	1007	63.0	2212	9	BC008317 Sequence
14	1007	63.0	2370	6	CQ834560 Sequence
15	1002	62.7	2815	10	BC054780 Sequence
16	959	60.0	1310	6	I76891 Sequence
17	959	60.0	2404	6	I76892 Sequence
18	875	54.7	1597	6	BD134432 Sequence
19	875	54.7	1597	6	AX017261 Sequence
20	875	54.7	1597	6	AX524965 Sequence
21	831	52.0	551	6	AX207216 Sequence
22	817	51.1	2229	5	AB126260 Sequence
23	740	46.3	1672	9	BC039498 Sequence
24	718	44.9	3309	6	CQ491400 Sequence
25	718	44.9	3309	6	CQ497275 Sequence
26	718	44.9	3760	9	BC073909 Sequence
27	718	44.9	4417	6	CQ414949 Sequence
28	718	44.9	5231	9	AB033091 Sequence
29	718	44.9	5322	6	AX405756 Sequence
30	712	44.5	3079	10	BC052880 Sequence
31	712	44.5	3865	10	BC059214 Sequence
32	712	44.5	5003	10	BC062918 Sequence
33	709	44.3	4395	10	AK122483 Sequence
34	674	42.2	3671	5	BC058056 Sequence
35	663	41.5	160170	2	AP001158 Sequence
36	659	41.2	221941	2	AC091060 Sequence
37	659	41.2	224788	2	AP001905 Sequence
38	634	39.6	736	6	CQ780369 Sequence
39	634	39.6	736	6	CQ780330 Sequence
40	634	39.6	736	6	BD125078 Sequence
41	634	39.6	736	6	BD125739 Sequence
42	634	39.6	2863	6	CQ782979 Sequence
43	634	39.6	2863	6	BD127410 Sequence
44	634	39.6	2863	9	AK074996 Sequence
45	594	37.1	152073	2	AC149703 Sequence

## ALIGNMENTS

RESULT 1	CQ834422	2744 bp	DNA	linear	PAT 29-JUL-2004
LOCUS	CQ834422	Sequence 293 from Patent WO2004058805.			
DEFINITION	CQ834422				
ACCESSION	CQ834422				
VERSION	CQ834422.1	GI:50833959			
KEYWORDS					
SOURCE					
ORGANISM	Homo sapiens (human)				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	Matsuda, A. and Yoneta, S.				
TITLE	T cell activating gene				
JOURNAL	Patent: WO 2004058805-A 293 15-JUL-2004;				
FEATURES	Asahi Kasei Pharma Corporation (JP)				
source	Location/Qualifiers				
	1..2744				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				
	138..2387				
	/note="unassigned protein product"				
	/codon_start=1				
	/protein_id="CAH05426.1"				
	/db_xref="GI:50833960"				
	/translation="MARLSTVILITFALSVTNPRLHKAAPQPTTEKISPMWESGI				
	MDLAISTRQYHQLQFLPYRGENNSLSVEGFRKLQNIQIDIKRIHIIHHDDHSHSH				

EHHSDEHSHSDHEHSHSDHEHSHSDHNHAAAGKRRKALCPDHSDSSCKDPRNSOGKCA  
 HRPEHAGRAVAKDSVASAEVTSYVTVTSVSEGTFFLEITIEPRGKLFPPDOVSTSP  
 SVTSKRSVRLAGKRNESVSEPRKQVNSNTNENPOECNASKLITSHMGQVPL  
 NATEPNYICPAIINOIDARSLIHTSEKKEIIPKTYSLQIAWGGIATISIFLSL  
 LGVILVLEPMRNVFPKFLISPLVALAVGSLSDAFIHLPHSHASHHSHHEPAMEM  
 KRGLPFLHLSQNI EESAYFDSTWKGTLALGGLFMPLVEVUTLILKQFKDKKRNOK  
 KPEMDDVEIKKOLSKYESQSLSTNEKVDTDRTGEGYLRADSOEPHFDSQPAVLEE  
 EEWMIAPHPQEVNVEYVPRCKNKKVSHFDITLQOSDULIHNNHDHNLHNNHON  
 HHPSHSORYSREELKDAGVATLAWMVMGDLMFSGDLIIGAAFTGELSSGLSTV  
 AVFCHELPHETGDFAVILKAGMTKQAVLYALASMLAYLGMATGIFIGHYAEVSWM  
 IFALTAGLPMYVALVDVYPERKLNDAHSDHSCSRNGYFPLQAGLILGFMILISIE  
 HKIVPRINE"

## ORIGIN

## Alignment Scores:

Score: 5,5e-112 Length: 2744  
 Percent Similarity: 1359,50 Matches: 301  
 Best Local Similarity: 40,32% Conservative: 1  
 Query Match: 40,19% Mismatches: 0  
 DB: 85,02% Indels: 447  
 Gaps: 1

US-10-659-004-104 (1-302) x CQ834422 (1-2744)

QY 1 MetalaarglyleuSerValleuileuthrPhealaleuSerValThrAsnPro 20  
 DB 138 ATGGGAGGAGAGTTATCTGTATCTTGATCTTGACCTTGGCCCTCTGTCAAAATCCC 197  
 QY 21 LeuHleGlyleuLyAlaAlaAlaPheProGlnThrGluLyLeuSesProAsnTrp 40  
 DB 198 CTTTCATGAACTTAAAGCAGCTGCTTCCCGACACACAGAGAAATTAATGTCGAAATGG 257  
 QY 41 GluSerGlyleuAsnValAspLeuAlaIleSerThrArgLysIleuGlnLeu 60  
 DB 258 GAAATCGCATTAATGTTGACTTGGCAATTTCCACCGCATATCATCTCAACAGCTT 317  
 QY 61 PheTyArgTyGlyGluAsnAsnSerleuSerValGluGlyPheArglyleuLeuGln 80  
 DB 318 TTCTACCGCTATGAGAAAATTAATCTTGTGATGAGGGGTTCAGAAATTAATCTTCA 377  
 QY 81 AenIleGlyleuAspLyseIleLyArgIleHileHleAspHisAspHisSer 100  
 DB 378 AATAATAGGATAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 437  
 QY 101 AspHisGluHisHisSerAspHisGluArgHisSerAspHisGluHisHisSerAsp 119  
 DB 438 GACCAAGAGATCATCATCAGATGAGCTCATCAGACCATGAGCATCATCATCAGACCATC 497  
 QY 119 ----- 119  
 DB 498 GAGCATCACTGACCATATCATCTGCTTCTGTAAATAAGGAAAGCTCTTGGC 557  
 QY 119 ----- 119  
 DB 558 CCAAGACCATGACTCAAGTAGTTAGSTAAAGATCTCTAGAAACGACGAGGGAAGAGCT 617  
 QY 119 ----- 119  
 DB 618 CACGACCAAGACATGCGAGTGTAGAGGAATGTCAAGACAGTGTAGTGTAGTAA 677  
 QY 119 ----- 119  
 DB 678 GTGACCTCACTGTGTACAACTGTCTGTGAAGAACTCACTTGTAGAGCAATAGAG 737  
 QY 119 ----- 119  
 DB 738 ACTTCAAGACCTGAAAACTCTTCCCAAGATGTAAAGAGCTCACTCCACCCACTGTC 797  
 QY 119 ----- 119  
 DB 798 ACATCAAGAGCGGGGTGAGCGGCTGTGTAGAGAAACAATGATCTGTAGTGTAG 857  
 QY 119 ----- 119

DB 858 CCGCGAAAGGCTTTATGTTATTTCCAGAAACAAATGAAATTCCTCAGAGTGTTCAT 917  
 QY 119 ----- 119  
 DB 918 GCATCAAGCTACTGACATCTCATGGCATGGGATCCAGTCCGGTGAATGCAACAGAG 977  
 QY 119 ----- 119  
 DB 978 TTCAACTATCTCTGTCCAGCATCATCAACCAATTGATGCTAGATCTTGTCTGATTCAT 1037  
 QY 119 ----- 119  
 DB 1038 ACAAGTAAAGAAAGGCTGAATCCCTCAAGACCTATTCATTACAAATAGCCTGGGT 1097  
 QY 119 ----- 119  
 DB 1098 GGTGTTTATAGCATTTTCCATCATCATGTTCTGTCTGTGGGGTTATCTTAGTG 1157  
 QY 119 ----- 119  
 DB 1158 CTTCTCATGAATCGGGTGTTCATTAATTTCTCTGAGTTCTTGTGGCACTGGCGGT 1217  
 QY 119 ----- 119  
 DB 1218 GGGACTTGAAGTGTATGCTTTTACACCTTCTTCCATTTCTCATGCAATGACACAC 1277  
 QY 119 ----- 119  
 DB 1278 CATATGATAGCATGAAGAACACGATGAATGAAGAGACCACTTTTCAATCAT 1337  
 QY 119 ----- 119  
 DB 1338 CTGTCTTCTCAAAACATGAAGAAAGTCTATTTTGTTCACAGTGAAGGTCTAACA 1397  
 QY 119 ----- 119  
 DB 1398 GCTTAGAGGCTGTATTTTATGTTCTTGTGAATGCTCTCATGATGATCAACAA 1457  
 QY 119 ----- 119  
 DB 1458 TTAAAGATTAAGAAAGAAAGATCAGAAAGCTGAAATGATGATGTGAGATT 1517  
 QY 119 ----- 119  
 DB 1518 AAGAAAGCTTGTCCAAATGATTAATCTCACTTTCAACAATGAGAGAAAGTATACA 1577  
 QY 119 ----- 119  
 DB 1578 GATGATGAGTGAAGGCTATTTACGAGCAGACTCACAAGAGCCTCCACTTGTATCT 1637  
 QY 119 ----- 119  
 DB 1638 CAGCAGCTGAGCTTGTGAAGAAAGAGGTGATGATGATGATGATGATGATGATGATG 1697  
 QY 119 ----- 119  
 DB 1698 GTTACAAATGATATGTACCCAGAGGTGCAAGATTAATGCAATTCATTTCCAGAT 1757  
 QY 119 ----- 119  
 DB 1758 ACATCGGCACTGACAGATCTATTCACCATCATGACTACATCATATTTCTCAT 1817  
 QY 120 ----- 119  
 DB 1818 CATCAACCAACCAAAACCATCTCTCAAGTCAACAGCAGCGCTACTCTGGAGAGAG 1877  
 QY 134 LeuLyAspAlaGlyValAlaThrLeuAlaTrpMetValIleMetGlyAspGlyLeuHis 153  
 DB 1878 CTGAAGATGCCGGGTGCGCACTTGGCTGAGATGATGATGATGATGATGATGATGATGATG 1937  
 QY 154 AenPheSerAspGlyLeuAlaIleGlyAlaAlaPheThrGluGlyLeuSerSerGlyLeu 173

Db 1938 AATTTCAGCGATGCGCTAGCAATTGCTGCTTTACTGTAAGGCTTATCAAGTGTTA 1997  
Qy 174 SerTherSerValAlaValAlaPheCysHisGluLeuProHisGluLeuGlyAspPheAlaVal 193  
Db 1998 AGTACTCTGTGTGCTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2057  
Qy 194 LeuLeuValAlaGlyMetThrValIleGlnAlaValLeuValLeuValLeuValLeuValLeuVal 213  
Db 2058 CTACTTAAGGCTGCGATGACCTTAAAGCAGGCTGCTCTTATTAATGCAATGCTGACCCATG 2117  
Qy 214 LeuAlaTyrlLeuGlyMetAlaThrGlyIlePheIleGlyHisGlyAlaGluAsnValSer 213  
Db 2118 CTGCGCATCTTGGATGGAATGCAACAGAAATTTTCAATGTCATTATGCGAAAATGTTTCT 2177  
Qy 234 MetTrpIlePheAlaLeuThrAlaGlyLeuPheMetHisValAlaLeuValAspMetVal 253  
Db 2178 ATGTGGATATTGCACTTACTGCTGCTTATTCATGATGATGCTGCTGCTGCTGCTGCTGCTGCTA 2237  
Qy 254 ProGluMetLeuHisAsnAspAlaSerAspHisGlyCysSerArgTrpGlyTrpPhePhe 273  
Db 2238 CTTAAATGCTGCAATGATGCTAGTGAACCAATGATGATGACCGCTGCGGATATTCTTT 2237  
Qy 274 LeuGlnAsnAlaGlyMetLeuLeuGlyPheGlyIleMetLeuLeuIleSerIlePheGly 293  
Db 2298 TTACAGATGCTGGAGTCTTTGGGTTTGGAAATTAATGTTACTTATTCCATATTGAA 2357  
Qy 294 HisValIleValPheArgIleAsnPhe 302  
Db 2358 CATAAATCGTGTTCGTATTAATTTTC 2384

RESULT 2  
LOCUS CO854099 2744 bp DNA linear PAT 23-AUG-2004  
DEFINITION Sequence 1 from Patent WO2004067564.  
ACCESSION CO854099  
VERSION CO854099.1 GI:51510129  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Homo sapiens  
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE  
AUTHORS Law, D., Gish, K.C., Murray, R. and Culp, P.  
TITLE Compositions against cancer antigen liv-1 and uses thereof  
JOURNAL Patent: WO 2004067564-A 1 12-AUG-2004;  
PROTEIN DESIGN LABS, INC. (US)  
FEATURES  
source 1..2744  
location/Qualifiers  
1..2744 /organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN

Alignment Scores:  
Pred. No.: 5-5e-112 Length: 2744  
Score: 1359.50 Matches: 301  
Percent Similarity: 40.32% Conservative: 1  
Best Local Similarity: 40.19% Mismatches: 0  
Query Match: 85.02% Indels: 447  
DB: 6 Gaps: 1

US-10-659-004-104 (1-302) x CO854099 (1-2744)

Qy 1 MetAlaArgIleLeuSerValIleLeuIleLeuThrPheAlaLeuSerValThrAsnPro 20  
Db 138 ATGCGAGGAATGATCTGTATCTGTATCTGTGACCTTGCCTCTGTCTGCAAAATCCC 197  
Qy 21 LeuHisGluLeuValAlaAlaPheProGlnThrThrGlyIleSerProAsnTrp 40  
Db 198 CTTATGAACTTAAAGCAGCTGCTTTCCCGCAGACCACTGAGAAATTAATGCGAATTTGG 257  
Qy 41 GluSerGlyIleAsnValAspLeuAlaIleSerThrArgGlnTrpHisLeuGlnGlnLeu 60

Db 258 GAATCTGGCATTAATTTGATCTTGCAATTTCACACGGCAATATCATCAACAGCTT 317  
Qy 61 PheTyrlArgTyrlGlyIleAsnAsnSerLeuSerValGlyIlePheArgIleLeuGln 80  
Db 318 TTCTACCGCTATGAGAAATAATTTCTTGTCACTTGAAGGTTCCAGAAATTTACTTCAA 377  
Qy 81 AsnIleGlyIleAspIleValArgIleHisIleHisAspHisAspHisHisSer 100  
Db 378 AATATGCGCATGATTAAGATTAAAGATTCATATACACCATGACACGACCATCACTCA 437  
Qy 101 AspHisGlyIleHisSerAspHisGlyIleArgHisSerAspHisGlyIleHisSerAsp 119  
Db 438 GACCAACGATCATCTGACATGATGAGGTACTCAGACCATGATCATCTCAGACAC 497  
Qy 119 ----- 119  
Db 498 GAGCATCATCTGACCATTAATCATGCTGCTTCTGTAAATAAGCGAAAGCTTTTGC 557  
Qy 119 ----- 119  
Db 558 CCAGACATGATCATGATGATGTTCAAGTAAGATCTTGAACACGCGGAAAGAGCT 617  
Qy 119 ----- 119  
Db 618 CACCGACCAAGACATGCCAGTGTAGAAAGAAATGCAAGACAGTGTAGTATGAA 677  
Qy 119 ----- 119  
Db 678 GTGACTCAACTGTGTACAACACTGTCTTGAAAGAACTCACTTCTAGACATAAGAG 737  
Qy 119 ----- 119  
Db 738 ACTCAAGACCTGGAAAACTTCTCCCAAGATGTAAAGAGCTCACTCCACCAAGTTC 797  
Qy 119 ----- 119  
Db 798 ACATCAAGAGCGGGTGAAGCGGCTGCTGTGTAGAAAAAATAATGATCTGTAGTAG 857  
Qy 119 ----- 119  
Db 858 CCGGAAAGGCTTTATGTATTCAGAAACAAATGAAATTCCTCAGAGTGTTCAT 917  
Qy 119 ----- 119  
Db 918 GCATCAAGCTACTGACATCTCATGATGATGGCATGCCAGTTCCGCTGAATGCAACAGAG 977  
Qy 119 ----- 119  
Db 978 TTCAACTATCTGTGCTCAGCATCATGACCAATTTGATGTATCTGTCTGATTCAT 1037  
Qy 119 ----- 119  
Db 1038 ACAATGAAAAAGAGCTGAATCCCTCAAAAGACTATTTCATTAACAATAGCTGGGT 1097  
Qy 119 ----- 119  
Db 1098 GGTGTTTATAGCATTTTCATCATGATGATTTCTGTCTGCTGGGGTTATCTTAGTG 1157  
Qy 119 ----- 119  
Db 1158 CCTCATGAATGGGGTGTTCATTAATTTCTCTGAGTTTCTTGTGGCATGGCGCTT 1217  
Qy 119 ----- 119  
Db 1218 GGGACTTTGAGTGTAGTCTTTTACACTTCTTCACATCTCATGCAAGTCAACAC 1277  
Qy 119 ----- 119  
Db 1278 CATATGATAGCATGAGAAACAGCATGAGAAATGAAAGAGACACTTTTCAATCAT 1337  
Qy 119 ----- 119  
Db 1338 CTGCTTCTCAAAATAGAAAGAGTGCATTTTGTATTCAGGTGAAAGGCTTACA 1397

[illegible]

ACCESSION	complete cds.
VERSION	U91060
KEYWORDS	U91060.2 GI:12711792
SOURCE	
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo. 1 (bases 1 to 2744)
AUTHORS	Green,C. and Morgan,H.
TITLE	Direct Submission
JOURNAL	Submitted (08-FEB-2001) Biochemistry, University of Liverpool, P.O. Box 147, Liverpool L69 3BX, UK
REFERENCE	2 (bases 1 to 2744)
AUTHORS	Green,C., Gillhool,E.M. and Walker,N.J.
TITLE	Direct Submission
JOURNAL	Submitted (21-NOV-1995) Biochemistry, University of Liverpool, P.O. Box 147, Liverpool L69 3BX, UK
COMMENT	On Feb 8, 2001 this sequence version replaced gi.1256000.
FEATURES	Location/Qualifiers
SOURCE	1..2744
gene	/organism="Homo sapiens"
CDS	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/cell_line="MCF-7 human breast cancer cell line"
	/note="estrogen induced mRNA"
	1..2744
	/gene="Liv-1"
	138..2387
	/gene="Liv-1"
	/note="breast cancer"
	/codon_start=1
	/product="estrogen regulated Liv-1 protein"
	/protein_id="AA96258.2"
	/db_xref="GI:12711793"
	/translation="MARKLSTVLITFLPALSVTNPLBKAAPQTEKISPNWESGGINVDALSTRQYHLQQLFYRGNNLSVEGRKLQIGIDIKRIIHNDHSHDHEHSDDHRSHSDHNSHDHNSDHNAAGNKAKLCIPRDSOGDPNDSOGKPEPSSGGTTPHREPHASGRNRVKDSVSASTVTSTNVTSSEGTFTLTETPRPGKKPKVSSSTPE SVTSKSVSRLAGRKINESVSEPKPMGTSNTENTENOECNASKLISHMGIOVPINATENLCPAIIINOIDARSLIHTSKKAIPEKXTSLQIAWGGLAISIFLSL LGVILIVDMNRVPERFLPSFVALAVGTLSGDAFLHPHSHAHHNHSHSEEPAMEXKRGLFSHLSNFIIEEASAFDSITWKGLTGALLGGLYFMFLVBHVTLIKQFKDKKKNO KRPNDVVETIKOLSKYESOLSSTEEKVDOTDRFGYLADSPSEHSDSOQPAVLEEEHWITAANHPOEVYNVEYVRGCCKKHSHPDITGOSDLITHHDVHNHLLHHNHON HHPHSQRYSREELKQAGVATTAMWTMDGILNFSDGLAIGAATFEGLSSGLSTS AVFHSELPHELGDFAVILLKAGMTVKQAVLYALSAMLAYLDMAVGTFIGHYAENVSMKI I FALTACLFMTVLVLDVWPENMLHNDASDHGCSRWGYFFLOAGMLLGIMLLISTPE HKIVERINF"
ORIGIN	
Alignment Scores:	
Pred. No.:	5.5e-112 Length: 2744
Percent:	1359.50 Matches: 301
Score Similarity:	40.32% Conservative: 1
Best Local Similarity:	40.19% Mismatches: 0
Query Match:	85.02% Indels: 447
DB:	Gaps: 1
US-10-659-004-104 (1-302) x HSUA1060 (1-2744)	
OY	1 MetaAlaaglygLeuSeValIlleleuIlleuthrPhealaleuSeVallthrAnPro 20
Db	138 ATGGCGAGGAAGTTATCTGTAATCTGTACTGACCTTGCCCTTCGTGCAAAATCCC 197
OY	21 LeuHlgileuLeValAlaAlapherProgInThrThrgulLyilLeserProanTrp 40
Db	138 CTTCATGAACTRAAAGCAGCTGCTTTCCCCCAGACCACTGAGAAATTAATGCCGAATGG 257
OY	41 GUseRgiylleasnValaspLeuAlalleSerTharGinTyrlHisleuglnleu 60
Db	258 GAATGTGGCATTAATGTTGACTTGGCAATTTCCACAGCAATATCATCTACAACAGCTT 317



KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 Schlegel, R., Endege, W.O. and Monahan, J.E.  
 AUTHORS Genes differentially expressed in human prostate cancer and their  
 TITLE use  
 JOURNAL Patent: WO 0160860-A 25736 23-AUG-2001;  
 Millennium Predictive Medicine, Inc. (US)

FEATURES  
 source  
 1. 3523  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 7,68e-112 Length: 3523  
 Score: 1359.50 Matches: 301  
 Percent Similarity: 40.32% Conservative: 1  
 Best Local Similarity: 40.19% Mismatches: 0  
 Query Match: 85.02% Indels: 447  
 DB: 6 Gaps: 1

US-10-659-004-104 (1-302) x CQ493869 (1-3523)

QY 1 MetAlARGLVLeuSerValIleLeuIleLeuThrPheAlaLeuSerValThrAsnPro 20  
 Db 138 ATGGCAGGAAATGTTATCTGTAATCTTGATCTGACCTTGGCCCTCTCTCAAAATCCC 197

QY 21 LeuHISGLuLeuLYaAlaAlaPheProGlnThrGluLYsIleSerProAsnTrp 40  
 Db 198 CTTGATGAACTAAAGCAGCTGCTTCCCGCAGCAGCAGAGAAATTAATGTCGAATTGG 257

QY 41 GluSerGlyIleAsnValAspLeuAlaIleSerThrArgGlnTyHisLeuGlnLeu 60  
 Db 258 GAATCGGATTAATGTTGATCTGGCAATTTCCACGGCAGATATCATCTTACACAGCTT 317

QY 61 PheTYArgTYrGlyGluAsnAsnSerLeuSerValGluGlyPheArgLYsLeuGln 80  
 Db 318 TTCTACCGCTAAGGAAATAATTTCTTGTGAGTTGAAGGGTTGAGAAATTAATTCTCA 377

QY 81 AsnIleGlyIleAspLYsIleLYsArgIleHisIleHisIleAspHisAspHisSer 100  
 Db 378 AATATAGGCAATGATTAAGATTAAAGATCCATATACCATGACACACCATCTCACTCA 437

QY 101 AspHisGluHisIleHisSerAspHisGluArgHisSerAspHisGluHisIleSerAsp 119  
 Db 438 GACCAAGAGCATCACTGACGACGACGTCCTCAAGACCATGACATCACTCAAGCCAC 497

QY 119 ----- 119

QY 498 GAGCATCACTGACCATATATCATGCTGTTCTGTAAATAAATAAGGAAAGCTTTTGC 557  
 Db 498 GAGCATCACTGACCATATATCATGCTGTTCTGTAAATAAATAAGGAAAGCTTTTGC 557

QY 119 ----- 119

QY 558 CCAAGACCATGATCAGATGATGCTAGTAAGATCTTAAGAAACAGCAGGGGAAAGAGCT 617  
 Db 558 CCAAGACCATGATCAGATGATGCTAGTAAGATCTTAAGAAACAGCAGGGGAAAGAGCT 617

QY 119 ----- 119

QY 618 CACCGACCAAGAACATGCGAGTGTAAAGAAATGTCAAGACAGCTGTAGTGTAA 677  
 Db 618 CACCGACCAAGAACATGCGAGTGTAAAGAAATGTCAAGACAGCTGTAGTGTAA 677

QY 119 ----- 119

QY 678 GTGACCTCAACTGTGTACAACACTGTCTGTGAAGAACTCACTTTCTAGAGACAATAGAG 737  
 Db 678 GTGACCTCAACTGTGTACAACACTGTCTGTGAAGAACTCACTTTCTAGAGACAATAGAG 737

QY 119 ----- 119

QY 738 ACTCCAGACCTGGAAAACTTTCCCAAGATGTAGAGACTCCACTCCACCAAGTGC 797  
 Db 738 ACTCCAGACCTGGAAAACTTTCCCAAGATGTAGAGACTCCACTCCACCAAGTGC 797

QY 119 ----- 119

Db 798 ACATCAAGAGCCGGTGAGCCGGCTGGTGTAGAGAAACAAATGAATCTGTAGTGAG 857  
 QY 119 ----- 119

Db 858 CCCCCAAAAGCTTTATGTATTTCCAGAAACAAATGAAAAATCTCAGAGTGTTCAT 917  
 QY 119 ----- 119

Db 918 GCATCAAGACTACTGACATCTCATGGCATGGGATCCAGGTTCCGCTGAATGCAACAGAG 977  
 QY 119 ----- 119

Db 978 TTCAACTATCTCTGTCCAGCATCATCAACCAATGTAGTATCTTGTGATTCAT 1037  
 QY 119 ----- 119

Db 1038 ACAAGTAAAGAGGCTGAATTCCTCAAGACCTATTCATTCAAAATAGCCTGGGTT 1097  
 QY 119 ----- 119

Db 1098 GGTGTTTATAGCCATTTTCATCATCATGTTCTCTGTCTGGGGTTATCTTAGTG 1157  
 QY 119 ----- 119

Db 1158 CCTCATGAATCGGGTGTTCCTCAAAATTTCTCCTGAGTTTCCTGTGGCACTGCGGTT 1217  
 QY 119 ----- 119

Db 1218 GGGACTTGTAGTGTATGCTTTTTCACCTTCTTCACATTCATGCAAGTCAACAC 1277  
 QY 119 ----- 119

Db 1278 CATAGCTATGACCATGAAGAACACGAATGGAATGAAGAGACCACTTTTCAGTCAT 1337  
 QY 119 ----- 119

Db 1338 CTGTCTTCAAAACATGAAGAAAGTGCTTATTTGATTCACGTGGAAGGTTCAACA 1397  
 QY 119 ----- 119

Db 1398 GCTTAGAGCCCTGTATTTTCATGTTTCTTGTGAACATGCTTCATTTGATCAAAACA 1457  
 QY 119 ----- 119

Db 1458 TTTAAAGTAAAGAAAGAAAGATCAGAAAGAACTGAAATGATGATGTGAGATT 1517  
 QY 119 ----- 119

Db 1518 AAGAGCAGTTGTCCAGTATGAATCTCACTTTCACAAATGAGAGAGAAAGTATACA 1577  
 QY 119 ----- 119

Db 1578 GATGATCGAAGTGAAGGCTATTTACGAGCAGACTCACAAGAGCCCTCCACTTTGATTCT 1637  
 QY 119 ----- 119

Db 1638 CAGCAGCTGAGTCTTGGAAGAGAGAGATCATGATAGTCAATGATCATCAGAGAA 1697  
 QY 119 ----- 119

Db 1698 GTCTCAATGAATATGTACCCAGAGGTGCAAGAAATGCAATTCACATTTCCAGAT 1757  
 QY 119 ----- 119

Db 1758 ACACTGGCCAGTCAAGCATCTTCAACCAATCATGACTACATCATATTTCTCAT 1817  
 QY 120 ----- 119

Db 1818 CATCAACCAACCAAAACACATCTCAGTCAAGCAGCGCTACTCTCGGAGGAG 1877  
 QY 134 LeuLYAspAlaGlyValAlaThrLeuAlaTrpMetValIleMetGlyAspGlyLeuHis 153

Dd		1878	CTGAAGAATGCCCGGTGGCCACTTGGCCTGAGTGAGTAATAAAGGATGAGCTGCAC	1937
Oy		154	AenpheserapgylyLeualaIilealyaaaphethrguglyleuserSerGyleu	173
Dd		1938	AATTTCAGCATGGCGCTAGCAATTGGTCTGTTTAACTGAAGCTTATCAAGTGTTTTA	1997
Oy		174	SerThrservAlAlavaPhecyeHieglueurProHieglueUglyAspPheAlAla	193
Dd		1998	AGTACTTCTGTGCTGCTTCTGTCATAGATTGCCCTCATGAATTAAGTGACTTTGCTGTT	2057
Oy		194	LeualeulyalaglyMetThrVallyeglnalaValeuryrAsnAlalaLeuserAlamet	213
Dd		2058	CTACTMAAGGCTGGCATACCGCTTAAAGAGGCTGCTCTTTAAATGCAITTGTCACCATTG	2117
Oy		214	LeualaTyreUglyMeclaAthrglyllepheileglsHtyrAlaglualsnValser	233
Dd		2118	CTGGGGTATCTTGGAAATGGACAAGAAATTTTCATGTGTCATTTAGCTGAATAAGTTCTT	2177
Oy		224	MettPllepheaIaleuthralaglyleupheuechIsvalAlaleuValaspmecylal	253
Dd		2178	ATGTGGAAATTTTGCATTACCTGCTGGCTTATTCAGTAAATGTTGCTGCTGGTGAATGTGTA	2237
Oy		254	ProglumetleuHisasnaspAlaseraphisaglyCysserArsgTrpglyTyrrPhephe	273
Dd		2238	CTGMAATGCTGCACATAAGTAGCTAGTACCACTGATGTACCGCTGGGGGATTTCTTT	2297
Oy		274	LeuglnaenalaglyMeclleuleuglylPheglyllemeleuleuleleserillepheglu	293
Dd		2298	TTCAGCAATGCTGGGATGCTTTTGGGCTTTTGGAAATTAATGTTACTTATTTCCATATTGCA	2357
Oy		294	HlelysllevalPheargIleasnDhe	302
Dd		2358	CATAAAATCGTGTTTCGTATAAATTTTC	2384

**RESULT 5**  
 AX465588  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 REFERENCE  
 AUTHORS  
 JOURNAL  
 FEATURES  
 SOURCE  
 CDS

AX465588	3461 bp	DNA	linear	PAT 16-JUL-2002
DEFINITION	Sequence 1 from Patent WO0216939.			
ACCESSION	AX465588			
VERSION	AX465588.1 GI:21899891			
KEYWORDS				
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Mack,D., Gish,K.C. and Wilson,K.E. Method of diagnosis of cancer and screening for cancer modulators Patent: WO 0216939-A 1 28-FEB-2002; EOS Biotechnology, Inc. (US) Location/Qualifiers 1..3461			
AUTHORS	/organism="Homo sapiens"			
JOURNAL	/mol_type="unassigned DNA"			
FEATURES	/db_xref="taxon:9606"			
SOURCE	/note="prostate/breast/ovarian/bladder cancer protein human BCR4 cDNA"			
CDS	138..2405 /note="unnamed protein product; human BCR4" /codon_start=1 /protein_id="CAD42374.1" /db_xref="GI:21899892" translation="MARKSLVILITLPAISVTNPDLHEAKAPPTOTEKISPMNESIGI NVDLAISTROYHLQOLFPRYRENNSSLSIEVGRKLONIGDKIKRINHNDHDHSDDPH EHSKHSHSDHSHENSHDHNSDDHNHHNHHAASKKRRALCPEDHSDSGSKPKPV SQGKAHRPEHASGRNRNVDSVASSEVTSTNTVSDEGTLELETERRPEKLPKQV SSSPVTSKSRVSRLAGRKTNESEBPRKGFMYSRTNENPOECFNASILTSHQGCM GIQVPLNATEFNVLCPAIINOIDARSCLIHTEKKAEIIRPTYSIGQTAWVGFTAISG ISPFLIGVILVPLMNRYEFERKLSFTVALAVGLTSGDAFILHPSHSHSHSHSB EPAMEMKEKRPFSHSSONIESAYFDSYTKGDIRLAGLYRMFYLVENVLTITIKPKXK KKKOYKRENDNDVEIKQLSKYESQULSTNEKQVTDTRDTGVYLRADSOESHSDDOQ PAVLEEVEVMIAAHAPVENVEYRGCKNCHSFHTLOQSDDLTHNDHYHIILH HHNHNNHPHSQRSYRSREELKDAGVATLAMVIMGDILNFSGLIALAFTEGLSS			

ORIGIN  
GLSTSVAVFCHLELPHLELGDFAVLLKAGTTVQAVLYNALSMLAYLGMAITQIFGTHZ  
ENVSMTMIFALTAGLEFMYVALVDVMTPEMLHNDASDHGCSRMGCFELQNAAGMLLGGIMD  
LISLFEKIVIRINF"

Alignment Scores:	
Pred. No.:	1,39e-111
Score:	1356.50
Percent Similarity:	40.00%
Best Local Similarity:	39.87%
Query Match:	84.83%
DB:	6
	Gaps: 1

US-10-659-004-104 (1-302) X AX465588 (1-3461)

QY	1	MetLaAaGlyLeuSerValIleLeuIleLeuThrPheAlaLeuSerValThrAspPro	20
Db	138	ATGGGAGGAAAGTTATCTGTATCTGTATCTGTACCTTGACCTTGGCTCTCTGTACAAATCCC	197
QY	21	LeuHISgluLeuLysAlaAlaAlaPheProGlnThrThrGluLysIleSerProAsnTrp	40
Db	198	CTTCAATGAACATAAAGCAGCTGGCTTCTCCCGACAGCACTGAGAAATTAAGTCCGATTGG	257
QY	41	GlusergIylIeAsnValAspLeuAlaIleSerThrArgIlnTrpHisLeuGlnLeu	60
Db	258	GAATCTGGCATTAATGTGTGACTTGGCAATTTCCACCGCAATATCATCTCAACAGCTT	317
QY	61	PheTYrArgTYrGluLysAsnAsnSerLeuSerValGluGlyPheArgLysLeuLeuGln	80
Db	318	TTCTACCCCTATGGAGAAATTAATCTTGTGTGACAGTGAAGGGTTCCAGAAATTACTTCA	377
QY	81	AsnIleGlyIleAspLysIleLysArgIleHisIleHisAspHisAspHisHisSer	100
Db	378	AATATAGGCATATGATTAAGATTAAAGATCCATATACCACTGACCAAGCATCATCTCA	437
QY	101	AspHisGluHisHisSerAspHisGluArgHisSerAspHisGluHisHisSerAsp---	119
Db	438	GACCAACGAGCATCACTCAGACCATGAGGCTCACTCAGACCATGAGCATCACTCAGACGAC	497
QY	119	-----	119
Db	498	GAGCATCACTGACCATGATCATCACTCTCACCATAATCATGCTGCTTCTGGTAAAT	557
QY	119	-----	119
Db	558	AAGGAAAAGCTCTTGGCCAGACCATGACTCAGATAGTTGACGTTAAAGATCTAGAATC	617
QY	119	-----	119
Db	618	AGCCAGGGGAAAGAGCTCACCGACAGAAACATGCCAGTGGTAGAAGAAATGTCAAGAC	677
QY	119	-----	119
Db	678	AGTGTAGTGCTAGTAAGTAGCTCACTGTGTACAACTGTCTCTGAAGAACTGAC	737
QY	119	-----	119
Db	738	TTTCTAGACAAATAGAGACTCAAGACTGTGAAAACTTTCCCAAAAGATGTAAAGACG	797
QY	119	-----	119
Db	798	TTCACCTCAACCAAGTGTACATCAAAAGAGCCGGGTGAGCCGGCTGGCTGTAGAGAAACA	857
QY	119	-----	119
Db	858	AATGATCTGTAGTAGAGCCCGGAAAAGCTTTATGTATTCAGAAAACAAATGAAAT	917
QY	119	-----	119
Db	918	CCTCAGAGGTGTTTCATATGATCAAAAGCTACTGACATCTCATGGCATGGGATCCAGGTT	977
QY	119	-----	119

```

Db      978 CCGCTGAATGCAAGAGTTCAATATCTCTGTCACGCCATCATCAACCAATTGATGCT 1037
Qy      119 -----
Db      1038 AGATTCTGTGATTCATACAGTGAAGAAAGAGCTGAATCCCTCAAGAGCTATTCA 1097
Qy      119 -----
Db      1098 TTACAAATAGCCTGGGTTGGTGTATTATAGCCATTTCATCATGATTTCTGTCTGTG 1157
Qy      119 -----
Db      1158 CTGGGGGTTATTTAGTGCCCTCATGATGGGGTGTTCCTCAATTTCTCTGAGTTTC 1217
Qy      119 -----
Db      1218 CTGTGGGACCTGGCCGTTGGGACCTTTGAGTGTGATGTTTTCACCTTCTTCACAT 1277
Qy      119 -----
Db      1278 TCTCATGCAAGTCACACCATATGATCATAGCCATGAAGAACAGCAATGGAATGAAGAA 1337
Qy      119 -----
Db      1338 GGACCACTTTTCAATCATCTGTCTCTCAAAACATGAAGAAAGTCCATTTTGATTC 1397
Qy      119 -----
Db      1398 ACGTGAAGGCTTAACAGCTCTAGAGAGGCTGTATTTTCATGTGTTTGTGAACATGTC 1457
Qy      119 -----
Db      1458 CTCACATTGATCAAACAATTTAAGTAAAGAAAGAAAGAAATCAGAGAAACCTGAAT 1517
Qy      119 -----
Db      1518 GATGATGATGTGAGATTAAAGAGAGTTGTCCAGTATGATCTCACTTTCACCAAT 1577
Qy      119 -----
Db      1578 GAGGAGAAAGTATGATACAGATGATGCACTGAAGGCTATTACAGAGAGCTCAGAGAG 1637
Qy      119 -----
Db      1638 CCTCCCACTTGTATTCTCAGAGAGCTGAGTCTTGAGAGAAAGAGTCAATGATGCT 1697
Qy      119 -----
Db      1698 CATGCTCATCCACAGAGAGTCAATGATATGTACCAAGAGGTGCAAGAAATAATGC 1757
Qy      119 -----
Db      1758 CATTCACATTTCCACGATACACTCGGCAGTCAGAGATCTCATTCACCACATCATGAC 1817
Qy      120 -----
Db      1818 TACCATCATATTTCTCATATATCCACCAACCAACCATCTTACAGTACAGCCAG 1877
Qy      128 ATGTYSerArgSgluLuleuLysAspAlaGlyValAlaThrLeuAlaTrpMetValIle 147
Db      1878 CGCTACTCTCGGAGAGAGCTGAAGATGCGGGCTGCCACTTTGGCCCTGAGTGGTGATA 1937
Qy      148 MetGlyAspGlyLeuHisAsnPhSerAspGlyLeuAlaIleGlyValAlaAphetrGlu 167
Db      1938 ATGGGATGAGCTGACCAATTTCAAGCATGCGCTGAGCAATGGTCTCTTTACTGAA 1997
Qy      168 GlyLeuSerSerGlyLeuSerThrSerValAlaValPheCysHisGlyLeuLeuProHisGlu 187
Db      1998 GGGTTATATCAAGGTGTTAAGTACTCTGTGTGTGTCTGTATAGTATGCTCTCATGAA 2057
Qy      188 LeuGlyAspPheAlaValLeuLeuLysAlaGlyMetThrValIleGlnAlaValLeuTyr 207
Db      2058 TTAGTGATCTTGTCTGTCTTACTTAAGGCTGSCATGACGTTAAGCAGGCTGTCTTAT 2117

```

```

Qy      208 AsnAlaLeuSerAlaMetLeuAlaTyrLeuGlyMetAlaThrGlyIlePheIleGlyHis 227
Db      2118 AATGCATTTGACAGCAGCTGGCGGTATCTTGAAATGGCAACAGAAATTTTCATGTCAT 2177
Qy      228 TyrAlaGlnAsnValSerMetTrpIlePheAlaLeuThrAlaGlyLeuPheMetHisVal 247
Db      2178 TATCTGAAATATGTTTCTATGTGATATTTGCACTTACTGTGCTTATTCATGATGTT 2237
Qy      248 AlaLeuValAspMetValProGluMetLeuHisAsnAspAlaSerAspHisGlyCysSer 267
Db      2238 GCTCTGTTGATATGATGTAACCTGAATGCTGACAAATGATGCTATGACCATGAGATGAGC 2297
Qy      268 ArgTrpGlyTyrPhePheLeuGlnAsnAlaGlyMetLeuLeuGlyPheGlyIleMetLeu 287
Db      2298 CGTGGGGGTATTTCTTTTACAGAAAGCTGGAGAGCTTTGGGTTTGGATTAATGTTTA 2357
Qy      288 LeuIleSerIlePheGluHisIleValPheArgIleAsnPh 302
Db      2258 CTTATTTCCATTTTGAACATAAATCGTGTTCGTATAATTTTC 2402

RESULT 6
LOCUS      AX829136          3461 bp      DNA      linear      PAT 12-DEC-2003
DEFINITION Sequence 29 from Patent WO02059377.
ACCESSION  AX829136
VERSION     AX829136.1  GI:39838917
KEYWORDS
SOURCE
  ORGANISM  Homo sapiens (human)
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
  1 Mack,D.H., Gish,K.C. and Afar,D.
    Methods of diagnosis of breast cancer, compositions and methods of
    screening for modulators of breast cancer
    Patent: WO 02059377-A 29 01-AUG-2002;
    EOS Biotechnology, Inc. (US)

FEATURES
  source
    1..3461
    /organism="Homo sapiens"
    /mol_type="unassigned DNA"
    /db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.:      1,39e-111      Length:      3461
Score:          1356.50      Matches:      301
Percent Similarity: 40.00%      Conservative: 1
Best Local Similarity: 84.87%      Mismatches: 0
Query Match:    84.83%      Indels:      453
DB:              6      Gaps:      1

US-10-659-004-104 (1-302) x AX829136 (1-3461)
Qy      1 MetAlaArgGlyLeuSerValIleLeuIleLeuThrPheAlaLeuSerValThrAsnPro 20
Db      138 ATGGGAGAGAGTATATCTGATCTTGAATCTTGAACCTTGGCCCTCTCTGTCACAAATCCC 197
Qy      21 LeuHisGluLeuLysAlaAlaAphetrGluThrIleGluLysIleSerProAsnTrp 40
Db      198 CTTCAATGAACATAAGAGAGCTGCTTCCCAAGACCACTGAGAAATTAAGTCGAATGG 257
Qy      41 GluSerGlyIleAsnValAspLeuAlaIleSerThrArgGlnTyrHisLeuGlnGlnLeu 60
Db      258 GAATCTGCATTAATGTTGACTTGGCAATTTCCACGCGCAATATCATCTACAAACGCTT 317
Qy      61 PheTyrArgTyrGlyGluAsnAsnSerLeuSerValIleGlyPheArgLysLeuGln 80
Db      318 TTCTACCGCTATGAGGAAATTAATTTCTTGTGAGTGAAGGTTCAAGAAATTAATTCTTCA 377
Qy      81 AsnIleGlyIleAspLysIleLysValArgIleHisIleHisAspHisAspHisSer 100

```

Dh 378 AATAAGGATAGATAGATTAAAGAAATTCATATACATGACGACGACATCATCA 437  
Qy 101 AepH1sgLH18H18serAepH18GLUH18serAepH18GLUH18H18serAepH18 119  
Dh 438 GACCACGACATCACTAGACCACTGAGCGTCACTCAGACCATGAGCATCACTAGACCA 497  
Qy 119 ----- 119  
Dh 498 GAGATCACTGACCATGATCATCTCTACCATTAATCATGCTCTTGTGTAATAAT 557  
Qy 119 ----- 119  
Dh 558 AAGGAAAAGCTTTTGCCAGACCATGATGATGATGATGATGATGATGATGATGAT 617  
Qy 119 ----- 119  
Dh 618 AGCCAGGGGAAAGAGCTCAACGACGAGAGATGCGAGTGAAGAAATGTCAGAGAC 677  
Qy 119 ----- 119  
Dh 678 AGTGTAGTGTGTAAGTGAAGTCACTCAAGTGTGAACAAGTGTGTGAAGAACTGAC 737  
Qy 119 ----- 119  
Dh 738 TTTCTAGAGACAATAGAGACTCAAGACCTGGAAAACTTTCCCAAGATGTAAGACG 797  
Qy 119 ----- 119  
Dh 798 TCCACTCCAGCAGTGTACATCAAAAGACCGGGTGAAGCGGCTGGTGAAGAAACA 857  
Qy 119 ----- 119  
Dh 858 AATGAATCTGTGAGTGAAGCGGCGGCTTTATGTATTCAGAAACAAATGAATAAT 917  
Qy 119 ----- 119  
Dh 918 CCTCAGAGTGTTCATGATCAAAAGCTAGATCTAGATGAGATGGGATCCAGGTT 977  
Qy 119 ----- 119  
Dh 978 CCGGTGAATGCAGAGAGTTCATCTGTCTGACGACATCATCAACCAATGATGCT 1037  
Qy 119 ----- 119  
Dh 1038 AGATCTGTCTGATTCATACAAAGGAAAGGCTGAATCCCTCCAAAGACTATTCA 1097  
Qy 119 ----- 119  
Dh 1098 TTACAAATAGCCTGGGTTGGTGTATATAGCATTTTCATCATGATTTCTGTCTG 1157  
Qy 119 ----- 119  
Dh 1158 CTGGGGGTTATCTAGTCTCTCATGATCGGGGTTTTCAAATTTCTCTGAGTTTC 1217  
Qy 119 ----- 119  
Dh 1218 CTGTGACATGCGCGCTTGGAATTGAGTGTGATGCTTTTACACCTTCTCCACAT 1277  
Qy 119 ----- 119  
Dh 1278 TCTCATGCAATGCACACCATAGTCATAGCCATGAGAACGACGAATGGAATAAGA 1337  
Qy 119 ----- 119  
Dh 1338 GACACACTTTTCAGTCATGTCTTCTCAAAAACATAGAGAAAGTCTTATTTGATTC 1397  
Qy 119 ----- 119  
Dh 1398 ACGTGAAGGCTTAACAGCTCTAGAGGCGCTGATTTTCATGTTTCTTTGGAACATGTC 1457  
Qy 119 ----- 119  
Dh 1458 CTCACATTGATCAACAATTTAAAGATTAAGAAAGAAATGACAGAAACCTGAATAAT 1517

Qy 119 ----- 119  
Dh 1518 GATGATGATGTGAGATTAGAGACAGTGTTCAGATGATCACTTCAACAAT 1577  
Qy 119 ----- 119  
Dh 1578 GAGAGAAAGTATGATACGATGATGATGATGATGATGATGATGATGATGATGAT 1637  
Qy 119 ----- 119  
Dh 1638 CCTCCCACTTTGATTTCTCAGACGCTGAGTCTTGAAAGAAAGAGGTGATGATGAT 1697  
Qy 119 ----- 119  
Dh 1698 CATGCTATCCACAGAAAGTCAATGATGATGATGATGATGATGATGATGATGATGAT 1757  
Qy 119 ----- 119  
Dh 1758 CATTCACTTTCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1817  
Qy 120 ----- 127  
Dh 1818 TACCATCATATTCTTCATCATCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1877  
Qy 128 ArgTySerArgGLUH18LeuLyAspAlaGLYAlaAlaThrLeuAlaTrpMetValIle 147  
Dh 1878 CGTACTCTCGGAGAGAGCTTAAAGATGCGCGCTGCGCACTTGGCTGTGATGTGATA 1937  
Qy 148 MetGlyAspGlyLeuH18AsnPheserAspGlyLeuAlaIleGlyAlaAlaIlePheThrGlu 167  
Dh 1938 ATGGGTGATGCGCTGACCAATTTTCAAGCATGCGCTTGAAGTGTGCTTTTACTGAA 1997  
Qy 168 GlyLeuSerSerGlyLeuSerThrSerValAlaValPheCysH18GluLeuProH18Glu 187  
Dh 1998 GCGTTATCAAGTGTGTTAAGTACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2057  
Qy 188 LeuGlyAspPheAlaValLeuLyAspAlaGLYAlaAlaThrLeuAlaTrpMetValIle 207  
Dh 2058 TTAGTGTACTTGTGTTCTACTAAAGCTGGCATGACCGTTAAGCAGGCTGTCTTTAT 2117  
Qy 208 AsnAlaLeuSerAlaMetLeuAlaTyrluGluH18MetAlaThrGlyIlePheH18GlyH18 227  
Dh 2118 AATGATTTGACGACAGCTGCGCTATCTTGAATGGCAACGAAATTTTCAATGCTCAT 2177  
Qy 228 TyrlaGluAsnValSerMetTrpIlePheAlaLeuThrAlaGlyLeuPheMetH18Val 247  
Dh 2178 TATGCTGAATAATGTTTCTATGATGATATTTGCACTTCTGCTGCTTATTCATGATGTT 2237  
Qy 248 AlaLeuValAspMetValProGluMetLeuH18AsnAspAlaSerAspH18GlyCysSer 267  
Dh 2238 GCTCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2297  
Qy 268 ArgTrpGlyTyrlPhePheLeuGluAsnAlaGlyMetLeuGlyIlePheGlyIleMetLeu 287  
Dh 2298 CGCTGGGGGTATTTCTTTTACAGAAAGCTGGAGAGCTTTTGGGTTTGGATTAATGTTA 2357  
Qy 288 LeuIleSerIlePheGluH18LeuAlaValPheArgIleAsnMetLeu 302  
Dh 2358 CTTATTTCCATATTGAAACATTAATCGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2402

RESULT 7  
AX511618  
LOCUS 3445 bp DNA linear PAT 27-SEP-2002  
DEFINITION Sequence 25 from Patent WO2055705.  
ACCESSION AX511618  
VERSION AX511618.1 GI:23392426  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE  
1

## AUTHORS

Mezes, P.S., Rastelli, L., Herrmann, J.L., MacDougall, J.R., Zhong, H.,  
Ceaman, S.J., Boldog, F., Shimkets, R.A., Gorman, L., Cresta, O.R.,  
Meyore, K.K., Folkerts, O., Martin, G.B., Eisen, A., Spaderna, S.K.,  
Vernier, C.A., Bergh, C., Spytek, K.A., DiIppio, V.A., Zernusen, B.D.,  
Peyman, J.A., Ellerman, K., Stone, D.J., Grose, W.M., Alsbrook, J.P.,  
Lejman, D.M., Rieger, D.K., Burgess, C.E. and Edinger, S.

## TITLE

Proteins and nucleic acids encoding same  
Patent: WO 02055705-A 25 18-JUL-2002;

## JOURNAL

Curagen Corporation (US)

## FEATURES

Location/Qualifiers  
1..3445  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

## ORIGIN

## Alignment Scores:

Pred. No.: 3,166-111 Length: 3445  
Score: 1352.50 Matches: 300  
Percent Similarity: 40.00% Conservative: 2  
Best Local Similarity: 39.74% Mismatches: 0  
Query Match: 84.58% Indels: 453  
DB: 6 Gaps: 1

US-10-659-004-104 (1-302) x AX511618 (1-3445)

QY 1 MetAlaArgLyLeuSerValIleLeuThrPheAlaLeuSerValThrAsnPro 20  
Db 117 ATGGAGAGAAATTATCTGTATCTTGATCTGACCTTGCCCTCTCTGTCAAAATCCC 176  
QY 21 LeuHISGLIuLeuValAlaAlaPheProGlnThrGluValIleSerProAsnTrp 40  
Db 177 CTTCATGAACTAAAGCAGCTGCTTCCCGACACAGAGAAATTAATTCGAAATGG 236  
QY 41 GluSerGlyIleAsnValAspLeuAlaIleSerThrArgGlnTrpHisLeuGlnIleu 60  
Db 237 GAATCTGGCATTAATGTTTACCTTGGCAATTCACACGCGCATATCATCTCAACAGCT 296  
QY 61 PheTyrArgTyrGlyGluAsnAsnSerLeuSerValGluGlyPheArgLyLeuLeuGln 80  
Db 297 TTCTACCGCTATGAGAAATTAATCTTGTCAATTGAAGGGTTCAGAAATTAATCTTCA 356  
QY 81 AsnIleGlyTyleAspLyValIleAspArgIleHisIleHisAspHisAspHisSer 100  
Db 357 AATATATAGCATATGATTAAGATTAAGAAATTCATATACCATGACCAACATCACTCA 416  
QY 101 AspHisGluHisHisSerAspHisGluArgHisSerAspHisGluHisHisSerAsp 119  
Db 417 GACCAAGAGATCACTCAAGCATGAGCGTCACTCAAGCATGAGCATCACTCAAGCAGC 476  
QY 119 ----- 119  
Db 477 GAGCATCACTGACCATGATCATCTCACTCAACATATATATGCTTCTGTAAATAAT 536  
QY 119 ----- 119  
Db 537 AAGCAAAAGCTCTTTGCCAGACATGACTGATGATAGTTCAAGTTAAAGTCTTAAAC 596  
QY 119 ----- 119  
Db 597 AGCCAGGGGAAGAGAGCTACGAGCAAGAAATGCGAGTGAAGAGATGCAAGAGAC 656  
QY 119 ----- 119  
Db 657 AGTGTAGTCTAGTGAAGTGAAGCTCAACTGTGTAGAACACTGTCTGTAAAGAACTCAG 716  
QY 119 ----- 119  
Db 717 TTTCTAGAGACAATAGAGACTCCAAAGCTTGAAAACTTTCCCAAAAGATGTAAGCAGC 776  
QY 119 ----- 119  
Db 777 TTCACATCCACCAAGTGTCAATCAAGAGAGCGGGGTGAGCGGGTGGTGTGTAGAAAACA 836

QY 119 ----- 119  
Db 837 AATGAATCTGTAGTAGAGCCCCGAGAAAGGCTTTATGTATTCAGAAAACAAATGAAAAAT 896  
QY 119 ----- 119  
Db 897 CCTCAGAGTGTTCATATGATCAAGAGTATCTGACATCTCATGAGCATGGGCATCAGGTT 956  
QY 119 ----- 119  
Db 957 CCGGTGAATGCAACAGAGTTCAACTATCTGTCCAGCATCATCAACAAATGATGCT 1016  
QY 119 ----- 119  
Db 1017 AGATCTGTGATTCATACAGTGAAGAGAGCTGAATCCCTCCAAAGACTATTCA 1076  
QY 119 ----- 119  
Db 1077 TTACAAATAGCCTGGGTGGTGGTTTATAGCCATTTCCATCATCAGTTTCTGTCTGTG 1136  
QY 119 ----- 119  
Db 1137 CTGGGGGTATCTTAGTCCCTCATGAAATCGGCTGTTTCAAAATTTCTCTGAGTTTC 1196  
QY 119 ----- 119  
Db 1197 CTGTGGCACTGGCCGTTGGGACTTTGAGTGTGATGCTTTTACACCTTCTTCACAT 1256  
QY 119 ----- 119  
Db 1257 TCTCATGAAGTACACACCATATGATCATAGCCATGAGAACGAGCAATGAAATGAAAAA 1316  
QY 119 ----- 119  
Db 1317 GAAACACTTTTACATCTGTCTTCAAAAATGAGAAAGTCCATTTGATTC 1376  
QY 119 ----- 119  
Db 1377 ACGTGAAGGATTAACAGCTCTAGAGGCGCTGATTTTCATGTTCTTGTGAACATGTC 1436  
QY 119 ----- 119  
Db 1437 CTCACATTGATCAAAATTAAGATTAAGAGAAAAAATCAGAGAAACCTGAAAT 1496  
QY 119 ----- 119  
Db 1497 GATGATGATGTGAGATTAAAGACAGATTGTCAAGTATGATCTCAACTTCAACAAT 1556  
QY 119 ----- 119  
Db 1557 GAGGAGAAAGTATGATACGATGATGCACTGAAGCTATTTACGAGCAGACTCAAGAG 1616  
QY 119 ----- 119  
Db 1617 CCGTCCACTTGTATTTCAAGAGCGTGCAGTCTTGAAGAAGAGAGTATGATAGCT 1676  
QY 119 ----- 119  
Db 1677 CATGCTCATCCAGAAAGTCTACATGAATATGATCCAGAGGATGCAAGAAATAATGC 1736  
QY 119 ----- 119  
Db 1737 CATTCACATTTCCAGATACACTGGCGCATGACATCTCATTCACACCATCATGAC 1796  
QY 120 ----- 127  
Db 1797 TACCATCATATTCTCATCATCAACCAACCAAAACCAACATCTTCAAGTCAAGCCAG 1856  
QY 128 ArgTyrSerArgGluGluLeuValAspAlaIleThrLeuAlaTrpMetValIle 147  
Db 1857 CGCTACTCTCGGAGAGAGCTGAAGATCGCGGCTGCGCATCTTGCGCTGGATGATGATA 1916



[illegible]

LOCUS	CG854111	2265 bp	DNA	linear	PAT 23-AUG-2004
DEFINITION	Sequence 13 from Patent WO2004067564.				
ACCESSION	CG854111				
VERSION	CG854111.1	GI:51510138			
KEYWORDS					
SOURCE	synthetic construct				
ORGANISM	synthetic construct				
REFERENCE	other sequences; artificial sequences.				
AUTHORS	1				
TITLE	Law, D., Gish, K. C., Murray, R. and Culp, P.				
JOURNAL	Compositions against cancer antigen liv-1 and uses thereof				
	Patent: WO 2004067564-A 13 12-AUG-2004;				
	PROTEIN DESIGN LABS, INC. (US)				
FEATURES	Location/Qualifiers				
source	1..2265				
	/organism="synthetic construct"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:32630"				
	/note="BCR4 mutant cDNA"				
ORIGIN					

Alignment Scores:	
Pred. No.:	5,86e-109
Score:	1324.50
Percent Similarity:	39.47%
Best Local Similarity:	39.34%
Query Match:	82.83%
DB:	6
Length:	2265
Matches:	297
Conservative:	1
Mismatches:	4
Indels:	453
Gaps:	1

QY	1	MetAlaArglyLeuSerValIlleuenuLeuThrPheAlaLeuSerValThirAspPro	20
Db	1	ATGGGAGAGAAAGTTATCTGTAACTTGATCTTGACTCTTGCCCTCTCTCTCTCAAAATGCC	60
QY	21	LeuHISgluLeuLysAlaAlaAlaPheProGlnThrGluLysIleSerProAsnTrp	40
Db	61	CTTCATGAAGCTAAAGACGCTGCTTTCGCCGACCACTGAGAAATATGTCCGAATTGG	120
QY	41	GluSerGlyIleAsnValAspLeuAlaIleSerThrArgGlnTyrHisIleuGIngluLeu	60
Db	121	GAATCTGGCATTAATGATGCACTTGGCAATTTCCACACGGCAATATCATCTTACAACGCTT	180
QY	61	PheTyrArgTyrGlyGluAsnAsnSerLeuSerValGluGlyPheArgLysLeuLeuGln	80
Db	181	TTCTACCCGCTATGAGGAAATATTAATCTTTGTGCTAGTTGAAGGTTTCAGAAATATCTTCAA	240
QY	81	AsnIleGlyIleAspLysIleLysArgIleHisIleHisAspHisAspHisIleSer	100
Db	241	AATATATGGCATGATGAATTAAGATTAAGATCCATATATACACATAGCACCAACCATCTCA	300
QY	101	AspHisGluHisHisSerAspHisGluTyrHisAspHisIleSerAspHisIleSerAsp	119
Db	301	GACCAAGACCATCACTCAGACCATGAGCCTCACTCAGACCAATGAGCATCACTTAGACCAAC	360
QY	119	-----	119
Db	361	GAGCATCACTTGACCATGATCATCTCACTCAACCATATCATCTGCTGTGTAATAAT	420
QY	119	-----	119
Db	421	AAGCGAAAAGCTTTTGCCAGACCATGATCAGATAGTTCAGGTAAAGATCCTAGAAC	480
QY	119	-----	119
Db	481	AGCCAGGGAAAGAGACTCAGCAGCAAGAACATGCCAGTGTAGAAAGAAATGTCAGAGAC	540
QY	119	-----	119
Db	541	AGGTGTAAGTCTAGTGAAGTCACTCACTGTGTACAACACTGTCTCTGAGGAACCTCAC	600
QY	119	-----	119

```

Db      601 TTCTAGAGACATAGAGACTCCAAGACCTGGAATACTTCCCAAGATGTAAGC 660
Qy      119 -----
Db      661 TCACCTCCACCCAGTGTACATCAAGAGCCGGGTGACCCGGCTGGGTAGAGAAAA 720
Qy      119 -----
Db      721 AATGATCTGTGAGTGAAGCCCGAAGAGCTTTATGTATTCAGAAAACAATGAAAT 780
Qy      119 -----
Db      781 CCTCAGAGTGTTCATATGATCAAGAGTACTGATCTCATGGATGGATCAGGTT 840
Qy      119 -----
Db      841 CCGGTGAATGCAAGAGTTCATATCTGTCTCAGCATCATCAACCAATTGATGCT 900
Qy      119 -----
Db      901 AGATCTTGTCTGATTCATACAGTGAAGAAAGGCTGAATCCCTCCAAAGACTATTCA 960
Qy      119 -----
Db      961 TTACAAATAGCCTGGGTTGGTGTATTATAGCAATTCATCATGATTTCTGTCTGTG 1020
Qy      119 -----
Db      1021 CTGGGGTTATCTTAGTGCTCTCATGAAATCGGGTGTTCCTGATTTCTCTGAGTTTC 1080
Qy      119 -----
Db      1081 CTGTGGCATGCGCGGTTGGGACTTTGAGTGTAGTCTTTTACACCTTTCTTCACAT 1140
Qy      119 -----
Db      1141 TCTCATGCAAGTCAACCATAGTCATAGCATGAGAAACAGCAATGGAATGAAAA 1200
Qy      119 -----
Db      1201 GGACCACTTTTCAGTCATCTGTCTTCATAAACAATAGAAAGAGTCTTATTTGATTC 1260
Qy      119 -----
Db      1261 ACGTGAAGGCTTAACAGCTCTAGAGAGGCTGTATTTTCATGTTCTTTGTAACATGTC 1320
Qy      119 -----
Db      1321 CTCACATTGATCAAAACAATTTAAGATAGAAGAAAGAAAGATCAGAAAGAACTGAAAT 1380
Qy      119 -----
Db      1381 GATGATGATGAGATTAAAGAGAGTGTCTCAAGTATGATCTCAACTTTCAAAAT 1440
Qy      119 -----
Db      1441 GAGAGAAAGTATGATACAGATGATGAACTGAAGGCTATTTACAGACAGACTCAAGAG 1500
Qy      119 -----
Db      1501 CCCCTCCCACTTGTATCTCAGACGCTGCACTTTGAGAAAGAGAGTATGATAGCT 1560
Qy      119 -----
Db      1561 CATCTCATCCAGAGAGTCTACATGATATGTACCCAGAGGTGCAAGATTAATGC 1620
Qy      119 -----
Db      1621 CATTCACATTTCCAGATACACTGGCCAGTCAAGCATTCATTCACCAACCATCATGAC 1680
Qy      120 -----
Db      1681 TACCATATATTCATCATACACCAACCAAAACCACTCTCTCAGAGTCAAGCAG 1740

```

```

Qy      128 ArgTyrSerArgGluGluLeuValAspAlaGlyValAlaThrLeuAlaTrpMetValIle 147
Db      1741 CGTACTCTCGGAGAGAGCTGAAGATGCGGCGCTCCCATTTGGCGCTGATGATGA 1800
Qy      148 MetGlyAspGlyLeuHisAsnPheSerAspGlyLeuAlaIleGlyValAlaPheThrGlu 167
Db      1801 ATGGGTATGAGCCCTGCACAATTTTCAGCGATGCGCTTACCAATGGTGTCTTACTGAA 1860
Qy      168 GlyLeuSerSerGlyLeuSerThrSerValAlaValPheCysHisGluLeuProHisGlu 187
Db      1861 GCGTTATCAAGTGTGTTAAGTACTTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCA 1920
Qy      188 LeuGlyAspPheAlaValIleLeuValAlaGlyMetThrValIleValAlaValLeuThr 207
Db      1921 TTAGGTACCTTGTGCTGTCTCTAATAAGCTGGCATGACCGTTAAGCAGAGTGTCTTAT 1980
Qy      208 AsnAlaLeuSerAlaMetLeuAlaTyrLeuGlyMetAlaThrGlyIlePheIleGlyHis 227
Db      1981 AATGATTTGTCAGCATGCTGCGCTATCTTGAAATGGCAACGAAATTTTCATTTGATCAT 2040
Qy      228 TyrAlaGluAsnValSerMetTrpIlePheAlaLeuThrAlaGlyLeuPheMetHisVal 247
Db      2041 TATGCTGAAATGTTTCTATGATGATATTGCACTTACTGCTGCTTATTCATGATGATGT 2100
Qy      248 AlaLeuValAspMetValProGluMetLeuHisAsnAspAlaSerAspHisGlyCysSer 267
Db      2101 GCTCTGTGTATTAAGTAAAGTCAAGATGCTGACATATGCTATGATGACATGATGATGAC 2160
Qy      268 ArgTrpGlyTyrPhePheLeuGluAsnAlaGlyMetLeuLeuGlyPheGlyIleMetLeu 287
Db      2161 GCGTGGGGTATTTCTTTTACAGAAATGCTGGATGCTTTGGGTTTGGAAATATGTTA 2220
Qy      288 LeuIleSerIlePheGluHisValIleValPheArgIleAsnPhe 302
Db      2221 CTATTTTCATATTTGAACATTAATCGTGTCTGTATTAATTTTC 2265

RESULT 10
AX207205 3461.bp DNA linear PAT 30-AUG-2001
LOCUS Sequence 1 from Patent W00155178
DEFINITION AX207205
ACCESSION AX207205
VERSION AX207205.1 GI:15394960
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Butheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1. Goddard, A., Gurney, A.L., Smith, V., Hongo, J.A. and de Sauvage, F.
   Compositions and methods for treatment of cancer
   Patent: WO 0155178-A 1 02-AUG-2001;
   GENENTECH, INC. (US)
FEATURES
    source
        1..3461
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 1,51e-107 Length: 3461
Score: 1311.50 Matches: 299
Percent Similarity: 40.05% Conservative: 1
Best local Similarity: 39.92% Mismatches: 2
Query Match: 82.02% Indels: 449
DB: Gaps: 1

US-10-659-004-104 (1-302) x AX207205 (1-3461)
Qy      1 MetAlaArgIleLeuSerValIleLeuIleuThrPheAlaLeuSerValThrAspPro 20
Db      138 ATGGCGAGAGAGTATCTGTATCTTATCTGATCTTGTGCTCTCTCTGATCAAAATCCC 197

```

QY 21 LeuHISgluLeuValaIaIaPheProGlnThrGluValIeSerProAntTp 40  
Db 198 CTTCAATGAATAAAGCAGCTGCTTCCCCAGACCACTGAGAAATATGATCCGATTTGG 257  
QY 41 GluSerGlyIleAsnValAspLeuValIeSerThrArgGlnTyrHisLeuGlnLeu 60  
Db 258 GAATCTGGCATTAAGTTGACTGGCAATTCACACGCAATATCATCTACAAACAGCTT 317  
QY 61 PheTyrArgTyrGlyGluAsnAsnSerLeuSerValGluGlyPheArgIleuLeuGln 80  
Db 318 TTCTACCCCTATGAGAAATATATCTTGTCAAGTTGAGAGGTTCAAGAAATTAAGTTCA 377  
QY 81 AsnIleGlyIleAspLysIleLysArgIleHisGlyHisHisAspHisAspHisSer 100  
Db 378 AATATAGGCATATGATTAAGATTAAAGAAATCCATATACACCATGACACGACCATCTACA 437  
QY 101 AspHisGluHisHisSerAspHisGluArgHisSerAspHisGluHisSerAsp 119  
Db 438 GACCGAGCATCTGACAGACCATGAGCGTCACTCAGACCATGACATCACTCAGACAC 497  
QY 119 ----- 119  
Db 498 GAGCATCACTGACCATATATCATGCTGCTTGTGTAATAATAAGCGAAAGCTTTGC 557  
QY 119 ----- 119  
Db 558 CCAGACCATGACTGATAGTTCAAGTAAGATCTTAAGAACAGCGAGGAAAGAGACT 617  
QY 119 ----- 119  
Db 618 CACCGACCAAGACATGCCAGTGTAGAAAGAAATGTCAGAGACATGTTAGTCTAGTAA 677  
QY 119 ----- 119  
Db 678 GTGACTCACTGTGTACAAACATGCTCTGTGAAGAACTCACTTCTAGAGACATAGAG 737  
QY 119 ----- 119  
Db 738 ACTCCAGACCTGGAATACTTTCCCAAGATGTAGAGAGCTCCACTCACCCAGTGTTC 797  
QY 119 ----- 119  
Db 798 ACATCAAGAGCCGGGTGAGCCGGCTGCTGTAGAGAAACAATGAATCTGTGAGTGAG 857  
QY 119 ----- 119  
Db 858 CCCCAGAAAAGCTTATGTATTCAGAAACAGAAATGAATAATCTCAGAGTGTTCAT 917  
QY 119 ----- 119  
Db 918 GCATCAAGACTATGACATCTCATGCGCATGGGCATCAGGTTCCGCTGAATCAACAGAG 977  
QY 119 ----- 119  
Db 978 TTCAACTATCTGTCTCAGCCATCATCAACAAATGATGTAGATCTTGTCTGATTCAT 1037  
QY 119 ----- 119  
Db 1038 ACAAGTGAAGAAAGGCTGAATCCCTCCAAAGACTATTCATTAACAATAGCTGGGTT 1097  
QY 119 ----- 119  
Db 1098 GGTGTTTTATAGCCATTTCCATCATCAGTTTCTGTCTCTGTGGGGTTATCTTAGTG 1157  
QY 119 ----- 119  
Db 1158 CCTCTCATGAATCGGGTGTTCCTCAATTTCTCTGAGTTCTCTGTGGGACTGGCGGTT 1217  
QY 119 ----- 119  
Db 1218 GGGACTTTGAGTGTGATGCTTTTTCACCTTTCTTCACATTTCTCATGCAAGTCAACAC 1277

QY 119 ----- 119  
Db 1278 CATATGATATGCCATTAAGAACCAAGCAATGGAATGAAGAGACCACTTTCAATCAT 1337  
QY 119 ----- 119  
Db 1338 CTGTCTTCTCAAAACATAGAAGAAAGTGCCTATTTGATTTCCAGTGGAAAGGTCTAACA 1397  
QY 119 ----- 119  
Db 1398 GCTCTAGAGGCGCTGATTTATCATGTTCTTGTGTAACATGCTCAGATTGATCAACAA 1457  
QY 119 ----- 119  
Db 1458 TTTAAAGATAGAAGAAAGAAAGATCAGAGAAACCTGAAATGATGATGTGAGATT 1517  
QY 119 ----- 119  
Db 1518 AAGAGCAGTTGTCCAGTATGATTCATCTCAACTTTCACAAATAGAGAGAAAGTAGATCA 1577  
QY 119 ----- 119  
Db 1578 GATGATCAACTGAAGCTATTTACGACAGACTCAGAGAGCCCTCCACTTGTATTCCT 1637  
QY 119 ----- 119  
Db 1638 CAGAGCCTGCACTTTGGAAGAAAGAGGTCAATGATGCTCATGCTCATCCACAGAA 1697  
QY 119 ----- 119  
Db 1698 GTTACATGAATATGTATCCAGAGGGGTCAAGAAATAATGCCATTCATTTCCAGCAT 1757  
QY 119 ----- 119  
Db 1758 AACTCGGCAGTACGATCTCATTCACCACCATCATGACTACCATATATCTTCAT 1817  
QY 120 ----- 133  
Db 1818 CATACACCAACCAAAACACCATCTCACAGTCAACCCGCTACTCTTCGGAGAGAG 1877  
QY 134 LeuLysAspAlaGlyValAlaThrLeuAlaThrMetValIleMetGlyAspGlyLeuHis 153  
Db 1878 CTGAAGATGCGCGGTGCCACTTTGGCTCGATGTGATATAGCGATGCGCTGCAC 1937  
QY 154 AsnPheSerAspGlyLeuAlaIleGlyAlaAlaPheThrGluGlyLeuSerSerGlyLeu 173  
Db 1938 AATTTCAAGCATGGCTGATGCAATTTGCTCTTTACTGAAGGCTTATCAAGTGTATA 1997  
QY 174 SerThrSerValAlaValPheCysHisGluLeuProHisGluLeuGlyAspPheAlaVal 193  
Db 1998 AGTACTTGTGTGCTGTGTCTGTCAATGATGCTCTCATGAATTAAGTGAATTTGCTGTT 2057  
QY 194 LeuLeuLysAlaGlyMetThrValLysGlnAlaValLeuTyrAsnAlaLeuSerAlaMet 213  
Db 2058 CTACTAAAGGCTGGATGACCGTTAAGCAGGCTGCTTTTAATGATGATTCAGCAGCATG 2117  
QY 214 LeuAlaTyrLeuGlyMetAlaThrGlyIlePheIleGlyHisTyrAlaGluAsnValSer 233  
Db 2118 CTGGCGTATCTTGGATGGCAACAGAAATTTTCATTTGATGCTTAAGCTGAAGAAATGTTCT 2177  
QY 234 MetTyrIlePheAlaLeuThrAlaGlyLeuPheMetHisValAlaLeuValAspMetVal 253  
Db 2178 ATGTGATATTTGCACTTACGCTGCTTATTCATGATGCTGCTGCTGCTGCTGCTGCTGCT 2237  
QY 254 ProGluMetLeuHisAsnAspAlaSerAspHisGlyCysSerArgTyrGlyTyrPhePhe 273  
Db 2238 CCTGAATGCTGCAACATGATGCTAGTACCATGATATAGCGCTGGGGGATTTTCTTT 2297  
QY 274 LeuGlnAsnAlaGlyMetLeuLeuGlyPheGlyIleMetLeuLeuIleSerIlePheGlu 293  
Db 2298 TTACAGATGCTGGGATGCTTTTGGTTTGGAAATTAATGTTACTTAT-TCCATATTTGAA 2356  
QY 294 HisLysIleValPheArgIleAsnPhe 302

DB	RESULT 11	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	PUBMED	REFERENCE	AUTHORS	TITLE	JOURNAL	REMARK	COMMENT
2357	CATATAATCGTCTCGT-ATAAATTC	2382																		
	BC055012	3610 bp	mRNA	linear	ROD 21-OCT-2003															
	BC055012		Mus musculus	solute carrier family 3 (metal ion transporter), complete cds.																
	BC055012		member 6, mRNA (cDNA clone MGC:62673 IMAGE:6337369), complete cds.																	
	BC055012.1	GI:32822908	MGC.																	
			Mus musculus (house mouse)																	
			Mus musculus																	
			Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.																	
			1 (bases 1 to 3610)																	
			Klausner R.D., Collins F.S., Wagner L.H., Shennan C.M., Schuler G.D.,																	
			Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,																	
			Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,																	
			Ditichenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,																	
			Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L.,																	
			Scerif T.E., Brownstein M.J., Udell T.B., Tothlyuk S.,																	
			Carnici P., Prange C., Raha S.S., Loquellano N.A., Peters G.J.,																	
			Abramson R.D., Mullany S.J., Bosak S.A., McSwan P.J.,																	
			McLernan K.J., Malek J.A., Gunaratne P.H., Richards S.,																	
			Worley K.C., Hale S.C., Garcia A.M., Gay L.J., Hulys S.W.,																	
			Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,																	
			Fahney J., Helton E., Kettelman M., Madan A., Rodriguez S.,																	
			Sanchez A., Whiting M., Madan A., Young A.C., Shcherenko Y.,																	
			Boulford G.C., Blakesley R.W., Touchman J.W., Green E.D.,																	
			Dickson G.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,																	
			Butterfield V.S., Krzywinski M.I., Skalka U., Smalls D.E.,																	
			Scherer A., Schein J.E., Jones S.J. and Marra M.A.																	
			human and mouse cDNA sequences																	
			Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)																	
			2 (bases 1 to 3610)																	
			Strausberg																	

/organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="MGC:62673 IMAGE:6337369"  
 /tissue\_type="Olfactory epithelium, neonatal mouse,  
 C57BL/6"  
 /clone\_1db="NIH\_MGC\_129"  
 /lab\_host="DH10B"  
 /note="vector: pCMV-SPORT6"  
 1..3610  
 /gene="Slc39a6"  
 /note="synonym: Emelin"  
 /db\_xref="LocusID:106957"  
 /db\_xref="MGI:2147279"  
 211..2508  
 /gene="Slc39a6"  
 /codon\_start=1  
 /product="Slc39a6 protein"  
 /protein\_id="AAH5012.1"  
 /db\_xref="GI:32822909"  
 /db\_xref="LocusID:106957"  
 /db\_xref="MGI:2147279"  
 /translation="MATDLSVIMLTLPALWTSPIHLEOSTAAPSOTTEKINSMBG  
 VNDVATYMQRHHIQQLPYRYGENDSLSYGPKKLQNTIGDIKRVHIIHDDHND  
 HEHSHDHHSHDHHSHDHHSHDHHSHDHHSHRSHSVAGKNNKATCPDSDSN  
 SGKQRTSLGKSGPAPAHMNGRNIKXSASSSEVTSAYVNVSEGRFVETITPKPG  
 RKNADVPSTPSTTEKSRVGRGLRSLARKKSNSVSEPKSFMYSRVTNDNIQECTP  
 TKLTSHGMSIQALNLTENPNYLCPAIINOIDRACLHFASEKKAIEPKYSLQIA  
 MGFQFIATISILSPILGTVILVPMNVPEPKILSPILVALGTSIGDALHILPSH  
 ASHSHSHSEEPANEMKRGPLFCHLSQNIESSYFDSTYKGLTALGGLYFMYLVEV  
 LELIKPQDKKKKKQCKPENDEDVESKQSLKSDSLSNEEVYDPEERESEYLRADS  
 QESPFPDQOPVLMLEEEVMIIAHNPQEVNVEYPRCKNKHSHFDITGQSDDLIH  
 HHNDYHIIILHHHHKQNHPPSHSORSEELKDAGIATLAMYIMGDGLNFDGLAI  
 GAATEGLSSGLSTSVAFQCHLPEHLEHGDPAVILKGMTVKQAVLYVALSAMLAYGDM  
 ATGFIGHYALENVSMVTFALTACLFMYVALVDMPENLHNDASDHGCSRNGYFLQNA  
 GILGFGIMLILISFEHKIVIRINE"  
 2011..2421  
 /gene="Slc39a6"  
 /note="ZIP; Region: ZIP zinc transporter. The ZIP family  
 consists of zinc transport proteins and many putative  
 metal transporters. The main contribution to this family  
 is from the Arabidopsis thaliana ZIP protein family these  
 proteins are responsible for zinc uptake in the plant.  
 Also found within this family are C. elegans proteins of  
 unknown function which are annotated as being similar to  
 human growth arrest inducible gene product, although this  
 protein is not found within this family"  
 /db\_xref="CDD:pfam02535"

ORIGIN

Alignment Scores:

Pred. No.:	Score:	Percent Similarity:	Best local Similarity:	DB:
6.92e-101	1237.50	38.43%	77.39%	10
Length:	Matches:	Conservative:	Mismatches:	Gaps:
3610	276	18	8	2

US-10-659-004-104 (1-302) x BC055012 (1-3610)  
 QY 1 Mel1a1a1rglyLeuSerVal1le1e1u1le1euThrPhe1a1a1e1uSerVal1trhAsnPro 20  
 Db 211 ATGGCAACAGATTATCTGTATATCATGATCTTGAACCTTTCGCCCTTGGATTACAGAGCC 270  
 QY 21 LeuH1sg1u1eu1y1a1a1--1A1a1a1PheProG1nThr1g1u1y1l1e1SerPro1a1n 39  
 Db 271 CTTCAATGAACATCAACATCAAGCGCTTCTCTCAGACTCTAGAGAAATTAATTCATAT 330  
 QY 40 Trp1u1se1rg1y1le1a1en1a1a1p1e1a1a1l1e1se1rh1r1a1rg1n1y1r1h1sl1e1u1g1n1 59  
 Db 331 TGGGAACCGTGTATTAAAGTTGACTTGG3CGATTACCAATGAGCAACCAATCTGCAGCAG 390  
 QY 60 LeuPhe1y1r1a1q1y1r1g1y1g1u1a1a1n1se1r1e1u1Se1r1e1u1a1g1u1y1Phe1a1r1y1Se1u1e1u 79

Db	391	CTATTTCACCGCTACCGAGAGAAAGATTCCCTTGCTCTGTAAAGCTTCAAAAAATTGGCTT	450
Qy	80	Glnaenilegilylleaplyeilleysargillehisilehisiaephi:saephi:shis	99
Db	451	CAGAACTATAGCATATGATATGATTTAAAGAGTCCATATACCATGACCAAGCATCAT	510
Qy	100	Seraphisgluhi:shis:Seraphisglu:arghis:Seraphisglu:hi:shis:Serap	119
Db	511	GCTGACCCAGCATCTACTCGACCATGAGCATCTCTGACCAAGCATCTCACTCGAC	570
Qy	119	-----	119
Db	571	CACGAGCATCACTGGACCAAGAGCATCTCGACCAAGACCAAGCATTTCCACCGAGT	630
Qy	119	-----	119
Db	631	CACACGGTTGCTGTGTAACAAATCGAAGACCTTTTGTCCAGACTTGACTGTATAT	690
Qy	119	-----	119
Db	691	TCAGGTAAAACTCTAGAACTAGTCTTAGGAAAGAGATCTCGCCACGACAGACATGAT	750
Qy	119	-----	119
Db	751	GGTAGAGAAACTTCAGAGAGAGTGCAGACTCTAGTAAAGTACCTGGCGTATACAC	810
Qy	119	-----	119
Db	811	GTTGTCTCTGAAGAACTCGCTTGTGTAGACAATAGAGACTCCAAAACTGGGAGAGC	870
Qy	119	-----	119
Db	871	ACCAAGATGTAAACCTTTCTTACCCACCCAGCATCAGAGAAAAAGCCGAGTGGCCG	930
Qy	119	-----	119
Db	931	CTGAGTCGGCTAGCTAGAGAAAAAGCATGAGTCTGTAGTGAAGCCAGAAAGCTTT	990
Qy	119	-----	119
Db	991	ATGTATTCAGAAACAAATGACAAATATAGAGTGTTCATATACACCAAGCTGCTG	1050
Qy	119	-----	119
Db	1051	ACATCCATGSGATGAGCATCCAGGCTCTGTGAATGCAAGGAATTTAATATCTGTC	1110
Qy	119	-----	119
Db	1111	CCAGCCATCATCAATCAATGATGCTCGGCTGTGTCTGATTCATACAGCAAGTAGAAG	1170
Qy	119	-----	119
Db	1171	AAAGCAGAAATCCCTCCAAAGACTATTTCTTTACAAATAGCTGCTGTGGCTTCATA	1230
Qy	119	-----	119
Db	1231	GCCATTTTCATATAGATTTCGTGCTGTGCTGAGGAGTCATCTTGTCACATGAAC	1290
Qy	119	-----	119
Db	1291	CGGATATTTTCAAGTTCTCTGTGAGCTTCTGTCGGCGCTGGCGCGGAACGCTGAGT	1350
Qy	119	-----	119
Db	1351	GGCGATCTCTGTACATCTTCTCCACACTCTCATGCAAGTCATCAGCAGATCATAGC	1410
Qy	119	-----	119
Db	1411	CATGAAGACGACGATGMAATGAAGAAGGCCCCCTGTCAGCACTGTGGCTCAG	1470
Qy	119	-----	119

LOCUS	AB071697	3287 bp	mRNA	linear	ROD 15-MAR-2002
Db	1471	AATATAGAGAAAGCTCCATATTTTGAATTCACAGTGGAAAGGTCTGACGGCTAGGGGGC	1530		
Qy	119	-----	119		
Db	1531	TTAATATTCATGTTCTGTGGAAACAGTACTCAGCTGATCAAGCAATTTAAGATAAG	1590		
Qy	119	-----	119		
Db	1591	AAAAAGAAAGATCAAAAAAGAACTGAAAATGATGAGAGTGTGAGAGCAAGAACGCTG	1650		
Qy	119	-----	119		
Db	1651	TCCAAATACGACTCTCAGCTTTCTCTAAATGAAGAGAGAGGTGACCCAGGGGAACGACT	1710		
Qy	119	-----	119		
Db	1711	GAAAGTATCTGCGAGCCGACTCCCAAGACCTCCCTCTTGTGATTTCCAGACGCGACG	1770		
Qy	119	-----	119		
Db	1771	ATGTTGGAAGAGAGAGGTGATGATAGCCCATGACACCCACAGAAAGTCTCAATGAA	1830		
Qy	119	-----	119		
Db	1831	TATGTGCCCAAGGGGCTGCAGAAACAAGTGCATTCACTTCCACGATACGTTGGCCAG	1890		
Qy	119	-----	119		
Db	1891	TCCGACGACCTCATCCACCAACCATCAACGACTACCATCATTCTGCACCAACCAAC	1950		
Qy	120	-----HSHISPROHISERHISERGLAARGTYSERARGUGLULYLAULYBAAPALA	137		
Db	1951	CAGAACCAACCACTCTCACAGCACAGGCAAGGCTACTCTTGAGAGGAGCTGAAAGAGCC	2010		
Qy	138	GLYVALAAlaThrLeuAlaIleTyrMetValIleMetGlyAspGlyLeuHisAsnSerAsp	157		
Db	2011	GGCATGTGCACATTGGCTGGAATGTGATCATGGGGGACGGGGCTGCACAAATTCAGTAGAC	2070		
Qy	158	GLYLeuAlaIleGlyAlaAlaIlePheThrGlyGlyLeuSerSerGlyLeuSerThrSerVal	177		
Db	2071	GGCTTGTGTTGGTGGTGGCTTCCACCGAGGGTTTGTCCAGTGGGCTAACCACTCTGTGC	2130		
Qy	178	AlaValAlaPheCysValIleGlyLeuPheProHisGlyLeuGlyAspPheAlaValLeuLeuValAla	197		
Db	2131	GCTGTGTTCTGTCAATGAACCTGCCCATGAACATAGTAATCTTGCTGTTTGGCTAAAGGCT	2190		
Qy	198	GlyMetThrValIleValGlnAlaValIleLeuTyrAsnAlaLeuSerAlaMetLeuAlaTyrLeu	217		
Db	2191	GGCATGACTGTCAACAGCAGCTGTCTCTATATCTCTGTACGCAAGTTGGGCTACCTT	2250		
Qy	218	GlyMetAlaThrGlyIlePheIleGlyHisIleTyrAlaGlyLeuAsnValIleSerMetTyrIlePhe	237		
Db	2251	GGAATGGCAACAGGGATATTCATCGGGCATATGCAAAAAATGTTTCTATGTGATATTC	2310		
Qy	228	AlaLeuThrAlaGlyLeuPheMetHisValAlaLeuValAlaPheMetValProGluMetLeu	257		
Db	2311	GCACCTACCTGCGGGTGTGTCATGATATGCTCTGTGTGACATAGGGCTCGAATGTG	2370		
Qy	258	HisAsnAspAlaSerAspHisGlyCysSerArgTyrGlyTyrPhePheLeuGlnAsnAla	277		
Db	2371	CACAAATGACTGATAGTACACAGATCAGCCGTTGGGGATATTTCTTCTCTCAGAAATGCT	2430		
Qy	278	GlyMetLeuLeuGlyPheGlyIleMetLeuLeuIleSerIlePheGluHisValIleVal	297		
Db	2431	GGGATATCTTCGCGTTTGGAAATATGTTACTCATTTCCATATTTGAGCATAAATTTGTG	2490		
Qy	298	PheArgIleAsnPhe302			
Db	2491	TTTCGATATAAATTC2505			

DEFINITION Mus musculus ermelin mRNA for endoplasmic reticulum membrane protein, complete cde.  
ACCESSION AB071697  
VERSION AB071697.1 GI:19570345  
KEYWORDS Mus musculus (house mouse)  
SOURCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
ORGANISM Mus musculus  
REFERENCE 1 Suzuki, A. and Endo, T.  
AUTHORS Ermelin, an endoplasmic reticulum transmembrane protein, contains the novel HELP domain conserved in eukaryotes  
TITLE Gene 284 (1-2), 31-40 (2002)  
JOURNAL MEDLINE  
PUBMED 21888618  
11891044  
REFERENCE 2 (bases 1 to 3287)  
AUTHORS Endo, T.  
TITLE Direct Submission  
JOURNAL Submitted (17-SEP-2001) Takeshi Endo, Chiba University, Department of Biology, Faculty of Science, 1-33 Yayoi-cho, Inage-ku, Chiba, Chiba 263-8522, Japan (E-mail: tendo@cuphd.nd.chiba-u.ac.jp, Tel:81-43-290-3911, Fax:81-43-290-3911)  
FEATURES  
source location/Qualifiers  
1..3287  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/cell\_line="C2"  
/cell\_type="skeletal muscle myoblasts"  
/clone\_id="lambda ZAPII C2 myoblast cDNA"  
1..3287  
/gene="ermelin"  
/gene="ermelin"  
704..2221  
/gene="ermelin"  
/codon\_start=1  
/product="endoplasmic reticulum membrane protein"  
/protein\_id="BAB86300.1"  
/db\_xref="GI:19570346"  
/translation="MYSRNTNDNIQECNTTKLITSHGSIQALLNATEFNYLCPAILNQIDPARACLIHTASGKAEIPKTYSLQIAMGFIATISIFSLGLGVILPLMNRVFFKFLSLVALAVGTLSDALHLPHSHASHSHSEPPAMKRGPLFSLSAQNISSSYFDSITWKGILTLGLYFMFLVHVTLIKOPDKKKKKKKKRENDVEDSKKQLSKYDSQLSNSNEKVDGERPESTLRADSESPFSDQPTMLEEERVTAAHPQEVNIEVPERGCKNCKSHFHDITGSDDLIHHDYHNLHHDYHNLHHDYHSHSQRYSRBEKDAGIATLAWVIMGDLNFSDLATGAATGEGISGLSTSVAFCHDELHELGDPAYILKAGMTVKQAVLYNALSAMLAYLGAATGIFIGYANVSMWIFALTAAGLFMYALVDMVPEMLHNDASDHGCSRMYGFLLGNAILLGFGLMISIFEHKIVRINF"  
ORIGIN  
Alignment Scores:  
Pred. No.: 6,496-93 Length: 3287  
Score: 1148.00 Matches: 260  
Percent Similarity: 37.35% Conservative: 16  
Best Local Similarity: 35.18% Mismatches: 6  
Query Match: 71.79% Indels: 458  
DB: 10 Gaps: 2  
US-10-659-004-104 (1-302) x AB071697 (1-3287)  
QY 21 leuhtagluuyluylala---alalaphproginthrrglulylleserproasn 39  
DB 3 cttatgaactacatcaacagctgcttctctcagctactgagaaaataatcaaat 62  
QY 40 ttrpsergylileasnvalaspleuallaleserthrrglulylrhielengln 59  
DB 63 tgggaacctggtttatcttgacttgccagcttaccatgcacgacacactctgcagc 122  
QY 60 leuherlyrargylrgylgluasnanserleuservalgluglyphearglyleu 79  
DB 123 ctattctaccctcctacgagagatgattcttgcctgttgaaagccttcagaaaattg 182  
QY 80 glmanilegylleaspleyleyleargliehliehliehliehliehliehlieh 99

DB 183 CAGAACTAGGCATATGATTAAGATCCATATACACATGACACAGACATCAT 242  
QY 100 SeraphisgluhihihiSeraphisgluarghihiSeraphisgluhihihiSerap 119  
DB 243 GCTGACACAGACATCACTCGGACCATATGATGATCTCGGACACAGACATCACTCGGA- 301  
QY 119 ----- 119  
DB 302 CCAGAGCATCATCGGACACAGACACATTCACCGCAGTCACACGTTGCTGTPA 361  
QY 119 ----- 119  
DB 362 AAACATGGAAGCCTTTTGTCCAGACTTGACTTGATTAATCAGSTAAAATCTTAC 421  
QY 119 ----- 119  
DB 422 AACTAGTGAAGAAAGATCTCGCCACAGACATGATGGTAGAGAACATCA 481  
QY 119 ----- 119  
DB 482 GGAGATGCAAGCTCTAGTGAAGTGAACCTCGCGGTATACAACGCTGTCTGAAGAAC 541  
QY 119 ----- 119  
DB 542 TCGCTTTAGAGCAATAGACTCCAAAACCTGGAGAGCACCAAGATGTAACT 601  
QY 119 ----- 119  
DB 602 TCTACCCACCCAGCATCAGGAGAAAGCCGAGTGGCCGCTGAGTCGCTAGTAGG 661  
QY 119 ----- 119  
DB 662 AAGAAAGCAATGAGTCTGTGAGTGAGCCAGAAAAGCTTATGTATTCAGAAACA 721  
QY 119 ----- 119  
DB 722 AATGACATATTCAGAGAGTGTTCATATCAACCAAGCTGTGACATCCCATGAGC 781  
QY 119 ----- 119  
DB 782 ATCCAGGCTGTGTAATGCAAGGAATTAATATCTCTGCCAGCCATCATCAATCA 841  
QY 119 ----- 119  
DB 842 ATTGATGCTCGGCTGTGTCTGATTCATACAGCAAGTAGAAGAAAGCAGAAATCCCTCA 901  
QY 119 ----- 119  
DB 902 AAGACTATTTCTTAACAATAGCCTGCTGTGCTTCAATAGCCATTTCCATCATCACT 961  
QY 119 ----- 119  
DB 962 TTCCTGTCTCTGAGAGTATCTGTGTGCACATGAACCGGTATTTTCAAGTTC 1021  
QY 119 ----- 119  
DB 1022 CTGCTGAGCTTCTCTGTGCGCCTGCGCTCGAAGCCTGATGCGCATCTCTGTACAT 1081  
QY 119 ----- 119  
DB 1082 CTTTCCCACTCTCATGCAAGTCATGACAGCATCATGACCATGAAGCCAGCGATG 1141  
QY 119 ----- 119  
DB 1142 GAATGAAGAAAGGCCCTTGTTCAGCCACCTGTCCGCTCGAATATAGAAAGATCC 1201  
QY 119 ----- 119  
DB 1202 TATTTGATTCACGTGAAAGTCTGACGCTCTAGGGGCTTATATTCATGTCTTCT 1261  
QY 119 ----- 119

Db 1262 GTGGAACAGTACTACACTGATCAAGCAATTTAAAGTAAGAAAAAGAAAGATCAAAAG 1321  
QY 119 ----- 119  
Db 1222 AAACCTGAAATGATGAGGATGTGGAGAGCAAGAACGACTGTCCAAATGACACTCTGAG 1381  
QY 119 ----- 119  
Db 1382 CTTTCTTCMAATGAGAGAGGTGAGCCAGGGGAAAGCACTGAAGGCTATCTGGAGACC 1441  
QY 119 ----- 119  
Db 1442 GACTCCCAAGAGCCCTCCCTTTGATTCCTCCAGACGCCGAGATGTGGAAAGAGAGAG 1501  
QY 119 ----- 119  
Db 1502 GTCATGATAGGCCCATGACACCCACAGAAAGTCTACATGAAATATGTGCCAGGGGCTGC 1561  
QY 119 ----- 119  
Db 1562 AAGAAAGATGCATTCACATTCCTCCAGATAGCTGGCCAGTCCGACGACTCTAC 1621  
QY 120 ----- 123  
Db 1622 CACCATCAGACTACCATCATTCTTGACACACACACACACACACACACACACACCTCAC 1681  
QY 124 SerHisSerGlnArgTyrSerArgGlnLeuLysValAlaPheCysHisGln 143  
Db 1682 AGCCACACGACGCTACTCTCGAGAGAGGCTGGAAGAGACGCGGATTCGCCACATTCGACC 1741  
QY 144 TTPMCTVALLIEMERGILYASPGILYENHISAENPESERASRGILYLEUALILEGLYLA 163  
Db 1742 TGGAGTGGATCATGATGAGGCGGCTGCACAAATTCAGTACGCGCTTGCTGATTTGCTGT 1801  
QY 164 AlaPheThrGlnGlyLeuSerSerGlyLeuSerThrSerValAlaValPheCysHisGln 183  
Db 1802 GCCTTCACCGAGGGTTTGTCCAGTGGGCTTAACACCTCTGTGCTGTCTGTATGAA 1861  
QY 184 LeuProHISGILYENGLYASRPHEALAVALLLEULYVALAGLYMETHTHVALLYSGIN 203  
Db 1862 CTGGCTCATGAACTAGTGAATCTTCTGCTGTTTGTAAAGGCTGGCATGCTGTCAAGCAG 1921  
QY 204 AlaValLeuTyrAspAlaIleuSerAlaMetLeuAlaTyrLeuGlyMetAlaThrGlyLe 223  
Db 1922 GCTGGCTTAAATGCTCTGTGACGCAATGTTGGCTTACTTGAATGGCAACAGAGATA 1981  
QY 224 PheileGlyHisTyrAlaGluAsnValSerMetTrrIlePheAlaLeuThrZlaGlyLeu 243  
Db 1982 TTCAATCGGGCATTAATGAGAAAATGTTTCTATGTGATTTTGCACCTCCTCCGGCTTG 2041  
QY 244 PheMetHisValAlaIleuValAspMetValProGluMetLeuHisAsnAspAlaSerAsp 263  
Db 2042 TTCATGATATGCTGCTGCTGTGATGATGAGGCTGAGATGTTCACACATGATCTGATGAT 2101  
QY 264 HisGlyCysSerArgTrrGlyTyrPhePheLeuGlnAsnAlaGlyMetLeuGlyPhe 283  
Db 2102 CACGAGATGACGCCGTTGGGATATTTCTTCCTGCAGAAATGCTGGGATATCTTCGGTTT 2161  
QY 284 GlyIleMetLeuLeuIleSerIlePheGlyHisIleValPheArgIleAsnPhe 302  
Db 2162 GGAATTAATTAATCTCATTTCCATATTGAGCATTAATAATTTGCTTTCGATTAATAATTT 2218

RESULT 13  
BC008317 2212 bp mRNA linear PRI 04-NOV-2003  
LOCUS Homo sapiens solute carrier family 39 (zinc transporter), member 6,  
DEFINITION mRNA (cDNA clone IMAGE:3343159), partial cds.  
ACCESSION BC008317  
VERSION BC008317.2 GI:34782897  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE  
AUTHORS  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
1 (bases 1 to 2212)  
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,  
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,  
Altschul,S.F., Zeeberg,B., Buettow,K.H., Schaefer,C.F., Bhat,N.K.,  
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,  
Diatchenko,L., Marusina,K., Farmer,A.A., Rabl,G.M., Hong,L.,  
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,  
Scheetz,T.E., Brownstein,M.J., Udell,T.B., Toshiyuki,S.,  
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,  
Abramson,R.D., Mulhaly,S.J., Bosak,S.A., McEwan,P.J.,  
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,  
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hu,Y.,S.W.,  
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,  
Faney,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,  
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,  
Bouffard,G.G., Blackesley,R.W., Touchman,J.W., Green,B.D.,  
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,  
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalhus,D.E.,  
Scherer,A., Schein,J.E., Jones,S.J. and Marra,M.A.  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
JOURNAL MEDLINE  
PUBMED 12477932  
2 (bases 1 to 2212)  
Strausberg,R.  
Direct Submission  
Submitted (25-MAY-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
On Sep 16, 2003, this sequence version replaced gi:14249878.  
Contact: MGC help desk  
Email: [cgapsb-remail.nih.gov](mailto:cgapsb-remail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Genome Sequence Centre,  
BC Cancer Agency, Vancouver, BC, Canada  
[info@bcgsc.bc.ca](mailto:info@bcgsc.bc.ca)  
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,  
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,  
Leticia Haiso, Martin Krzywinski, Beta Kutsche, Oliver Lee, Soo  
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeary, Steven  
Neas, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Seedei, Jacqueline  
Schein, Duane Smalhus, Michael Smith, Ioraine Spence, Jeff Stott,  
Michael Thorne, Miranda Teal, Nataesja van den Bosch, Jill Vardy,  
George Yang, Scott Zuyderduyn, Marco Marra.  
Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IPAL Plate: 22 Row: K Column: 4.  
Location/Qualifiers  
1..2212  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3343159"  
/issue\_type="Eye, retinoblastoma"  
/clone\_id="NIH MGC 16"  
/lab\_host="DH10B-R"  
/note="Vector: pOTB7"  
1..2212  
/gene="SLC39A6"  
/note="synonym: LIV-1"  
/db\_xref="locusID:25800"  
1..1127  
/gene="SLC39A6"  
/codon\_start=3  
/product="SLC39A6 protein"  
/protein\_id="AAH08317.2"



OY		144	TTPMevall1enwctg1yaerpjyleuh1asnpsheserbpjyleua1alleg1yala	163
Db		830	TGGATGTGTAATAAGGGGTGATGCGCTGCACATTTCAGCAGTGGACTGAATTCGGTGCT	889
OY		164	Alaphethrgug1yleusersegc1yleuserthiservala1avalphecysahisglu	183
Db		890	GCTTTTACTGAGGCTTATCAAGTGGTTAAGTACTTCTGTGGTGTGTTGTGCATGAG	949
OY		184	LeudProhisgu1eug1yaspPheal1avalleuleulyalaglymeethvallyglu	203
Db		950	TTGCCCATGAAATTAGGTGACTTTGCTGTTCTACTTAAGGCTGCATACCGTTAAGAG	1009
OY		204	AlaValleu1yrana1aleuser1awmetleula1tytleug1ymeFalathrgh1yle	223
Db		1010	GCTGTCCCTTAATKAANTGCAATTCACACCAGTCTGGCGTATCTTGGAAATGGACAAGAAAT	1069
OY		224	Pheilegiyh1st1ryZALaGLuaSnValsemctTpilipehalaleuth1Alag1yeu	243
Db		1070	TTCAATGGTCAATTAATGCTGAAAAGTCTTCTAATGTGATATTGGACTTACTGCTGGCTTA	1129
OY		244	PheMeCh1SvAlAlaLeuValAspMetValProglumetleuH1SaPaAspASerSp	263
Db		1130	TTCAATGATGTGGCTCTGTGTGATGTGTACCTGAAGAAGCTGCACAAGAATGCTAGTAC	1189
OY		264	H1eg1CyasearXTTg1LYTyZpHePheLeuGLuaAnaGL1MeLeuLeng1yPhe	283
Db		1190	CATGATATGACCGCTGGGGGATTTCTTTTACAGAAATGCTGGGATGCTTTGGGTTTT	1249
OY		284	G1yl1emettLeuLeu1esertllephegunih1syll1eval1Phearg11easnPh	302
Db		1250	GGAATTATGTTACTTATTCCATATTGGAACATAAATCGTGTTCGTATAAATTTTC	1306
RESULT 15				
BC054780		2815 bp	mRNA	linear
LOCUS	Mus musculus endoplasmic reticulum membrane protein, mRNA (CDNA			
DEFINITION	clone IMAGE:681386), partial cds.			
ACCESSION	BC054780			
VERSION	BC054780.1 GI:32493411			
KEYWORDS				
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus,			
	1 (bases 1 to 2815)			
	Straussberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,			
	Klausner,R.D., Collins,F.S., Wagner,L., Shennan,C.M., Schuler,G.D.,			
	Altshuler,S.F., Zeeberg,B., Buetow,K.H., Scheffer,C.F., Bat,N.K.,			
	Hopkins,R.P., Jordan,H., Moore,T., Max,A.I., Wang,J., Haleh,F.,			
	Ditachenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,			
	Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,			
	Scheetz,T.E., Brownstein,M.T., Ueda,T.B., Toshiyuki,S.,			
	Carinci,P., Prange,C., Kahn,S.G., Loggellano,N.A., Peters,G.J.,			
	Abriemson,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.D.,			
	McKenna,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,			
	Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Huijck,S.W.,			
	Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,			
	Rahcy,U., Helton,E., Keteleman,M., Madan,A., Rodrigues,S.,			
	Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shenkevich,Y.,			
	Bouffard,G.G., Blakesley,R.W., Touchman,D.W., Green,E.D.,			
	Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schultz,J., Myers,R.M.,			
	Butterfield,Y.S., Krzywinski,M.I., Skalski,U., Smallus,D.,			
	Scherch,A., Schein,J.E., Jones,S.J. and Marra,M.A.			
TITLE	Generation and initial analysis of more than 15,000 full-length			
JOURNAL MEDLINE	human and mouse cDNA sequences			
PUBMED	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)			
REFERENCE	22388257			
AUTHORS	2 (bases 1 to 2815)			
TITLE	Straussberg,R.			
JOURNAL	Submitted (01-JUL-2003) National Institutes of Health, Mammalian			

[illegible]

```
Qy 112 SerAspHisGluHisSerAspHisProHisSerHisSerGlnArgTyrSerArg 131
Db 1182 ----CACCAACACACACAGAACCAACCACTCCACAGCCACAGCCAGGCTACTCTCGA 1235
Qy 1332 GluGluLeuLysAspAlaGlyValAlaThrLeuAlaTrpMetValIleMetGlyAspGly 151
Db 1236 GAGGAGCTGAAGAGAGCCGGCATTCGCAATTGGCCTGGATGGTGCATCATGGGCGACGGG 1295
Qy 152 LeuHisAsnPheSerAspGlyLeuAlaIleGlyValAlaPheThrGluGlyLeuSerSer 171
Db 1296 CTGCACAAATTCAGTGAAGGCGCTGCTATGGTCTGCTTCACCGAGGGTTTGTCCAGT 1355
Qy 172 GlyLeuSerThrSerValAlaValPheCysHisGluLeuProHisGluLeuGlyAspPhe 191
Db 1356 GGGCTAAGCACTCTGTGCTGTGTTCTGTCTCATGAACTGCTCATGAACCTAAGTGACTTT 1415
Qy 192 AlaValLeuLeuLysAlaGlyMetThrValLysGlnAlaValLeuTyrAsnAlaLeuSer 211
Db 1416 GCTGTTTGGCTAAAGGCTGGCATGACTGTCAAGCAGGCTGTGCTATATATGCTCTGTCA 1475
Qy 212 AlaMetLeuAlaTyrLeuGlyMetAlaThrGlyIlePheIleGlyHisTyrAlaGluAsn 231
Db 1476 GCCATGTTGGCTTACCTTGGAAATGCAACAGGATATTCATCGGCGCATTAATGCAAGAAAT 1535
Qy 232 ValSerMetTrpIlePheAlaLeuThrAlaGlyLeuPheMetHisValAlaLeuValAsp 251
Db 1536 GTTCTATGTGGAATATTCGACTGCTGCGGCTGTTCATGTATGTGCTGTGTTGAC 1595
Qy 252 MetValProGluMetLeuHisAsnAspAlaSerAspHisGlyCysSerArgTrpGlyTyr 271
Db 1596 ATGGTGCTGAGATGTTGCACAAATGATGTAATGATGATGATGATGATGATGATGATGAT 1655
Qy 272 PhePheLeuGlnAsnAlaGlyMetLeuLeuGlyPheGlyIleMetLeuLeuIleSerIle 291
Db 1656 TTCTTCCTGCGAATGCTGGGATGCTTCTCGGTTTGGAAATATGTTACTCATTTCCATA 1715
Qy 292 PheGluHisLysIleValPheArgIleAsnPhe 302
Db 1716 TTGAGCATTAATAATTGTGTGTAATTAATTTTC 1748
```

Search completed: March 20, 2005, 00:33:45  
Job time : 4947 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

# OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 19, 2005, 20:53:38 ; Search time 617 Seconds  
(without alignments)  
2897.507 Million cell updates/sec

Title: US-10-659-004-104

Perfect score: 1599  
Sequence: 1 MARKUSVILITLFPALSVNP.....FGIMLLISIFHKIVFRINF 302

## Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Command line parameters:

-MODEL=frame+\_p2n\_model -DEV=x1h  
-O=/cgn2.1/USPRO/seq01/US10659004/runat\_14032005\_181234\_15731/app\_query.fasta\_1.455  
-DB=N\_Geneseq\_16Dec04 -QFMT=fastcap -SUFFIX=ring -MINMATCH=0.1 -LOOBL=0  
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blowum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODS=LOCAL -OUTFMT=pro -NOR=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10659004 @CGN\_1\_1\_470 @runat\_14032005\_181234\_15731 -NCPU=6 -ICPU=3  
-NO MMAR -LARGEOQUERY -NEG\_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOC  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAEXT=7 -YGAPOP=10 -YGAEXT=0.5 -DELOP=6 -DELEXT=7

## Database :

N\_Geneseq\_16Dec04:\*  
1: geneseqn1980s:\*  
2: geneseqn190s:\*  
3: geneseqn200s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002as:\*  
7: geneseqn2002bs:\*  
8: geneseqn2003as:\*  
9: geneseqn2003bs:\*  
10: geneseqn2003cs:\*  
11: geneseqn2003ds:\*  
12: geneseqn2004as:\*  
13: geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1599	100.0	1101	12	ADL57158
2	1359.5	85.0	2250	13	ADR28033
3	1359.5	85.0	2744	8	ACC50210
4	1359.5	85.0	2744	10	ADH28902
5	1359.5	85.0	2744	12	ADN03826

6	1359.5	85.0	2744	12	ADQ96115	Adq96115 T cell ac
7	1359.5	85.0	2744	13	ADR25539	Adr25539 Breast ca
8	1359.5	85.0	2744	13	ADR87269	Adr87269 CDNA enco
9	1359.5	85.0	2744	13	ADP56127	Adp56127 Human pro
10	1359.5	85.0	2744	13	ADR66497	Adr66497 Human pro
11	1359.5	85.0	2744	13	ADR66155	Adr66155 Human pro
12	1359.5	85.0	3523	5	ABV25747	Abv25747 Human pro
13	1359.5	85.0	3523	11	ACN88857	Acn88857 Breast ca
14	1359.5	84.8	2268	8	ACC51044	Acc51044 Human dia
15	1356.5	84.8	2268	12	ADL57160	Adl57160 Human NOV
16	1356.5	84.8	3461	6	ABR92205	AbR92205 Prostate
17	1356.5	84.8	3461	6	ABT07707	AbT07707 Breast ca
18	1356.5	84.8	3461	6	ABA92299	AbA92299 Human bre
19	1356.5	84.8	3461	8	ABX76337	AbX76337 Lung can
20	1356.5	84.8	3461	11	ADN38737	Adn38737 Cancer/an
21	1356.5	84.8	3461	11	ADN39560	Adn39560 Cancer/an
22	1356.5	84.8	3461	11	ADN39520	Adn39520 Cancer/an
23	1355	84.7	1186	12	ADL57156	Adl57156 Human NOV
24	1352.5	84.6	2763	13	ACN40337	Acn40337 Tumour-as
25	1352.5	84.6	3445	6	ABS59534	AbS59534 Human CDN
26	1352.5	84.6	3445	12	ADH72211	Adh72211 Human gen
27	1352.5	84.6	3445	12	ADL57152	Adl57152 Human NOV
28	1351.5	84.5	2310	12	ADL57162	Adl57162 Human NOV
29	1349.5	84.4	2776	4	ADL13480	Adl13480 Human LIV
30	1347.5	84.3	3445	12	ADL57168	Adl57168 Human NOV
31	1345.5	84.1	3445	12	ADL57166	Adl57166 Human NOV
32	1345.5	84.1	3445	12	ADO39199	Ado39199 Human CDN
33	1340.5	83.8	3445	12	ADL56815	Adl56815 Human NOV
34	1324.5	82.8	2265	13	ADR87281	Adr87281 Liv-1 mut
35	1311.5	82.0	3461	4	ADL13479	Adl13479 Human LIV
36	1311.5	82.0	3586	10	ABT31919	Abt31919 Human bre
37	1265.5	79.1	2211	12	ADL57164	Adl57164 Human NOV
38	1189.5	74.4	3173	13	ACN43699	Acn43699 Human dia
39	1007	63.0	2236	12	ADN03825	Adn03825 Antipsoi
40	1007	63.0	2370	12	ADQ96253	Adq96253 T cell ac
41	959	60.0	1310	2	AAT33219	Aat33219 Oestrogen
42	959	60.0	1310	2	AAT99070	Aat99070 Partial s
43	959	60.0	2404	2	AAT33220	Aat33220 Oestrogen
44	959	60.0	2404	2	AAT99071	Aat99071 Partial s
45	875	54.7	1597	2	AAZ33622	Aaz33622 Human bre

## ALIGNMENTS

RESULT 1	ADL57158	ADL57158 standard; DNA; 1101 BP.
XX	ADL57158	
XX	ADL57158	
XX	ADL57158	
DT	03-JUN-2004	(first entry)
XX		
DE	Human NOV9c	gene SEQ ID NO:103.
XX		
KW	ds, gene; human; antidiabetic; anorectic; cardiac; hypotensive;	
KW	antiartherosclerotic; anorectic; virucide; antibacterial; fungicide;	
KW	prolactinase; neurotrophic; neuroprotective; antiparthenon;	
KW	anticonvulsant; osteopathic; antiarthritic; antiinflammatory;	
KW	dermatological; antiasclerotic; antidiabetic; gene therapy;	
KW	fibroblast growth factor receptor 4; FGFR4;	
KW	complement factor I precursor; matrix metalloproteinase-15 precursor;	
KW	MDG3; T-lymphocyte surface antigen Ly-9 precursor;	
KW	fibroblast growth factor-21; FGF-21;	
KW	alpha-2 macroglobulin-like polypeptide variant;	
KW	antileukoproteinase 1 precursor; LIV-1; nuclear hormone receptor NOR-1;	
KW	transmembrane protein-like; beta-neoendorphin-dynorphin precursor.	
OS	Homo sapiens.	
XX		
Key	Location/Qualifiers	
PH	123..1031	
CDS		
FT	/+tag= a	
XX		

PN WO2004022723-A2.  
 XX 18-MAR-2004.  
 PD 09-SEP-2003; 2003WO-US028141.  
 XX 09-SEP-2003; 2002US-0409145P.  
 PR 10-SEP-2002; 2002US-0409544P.  
 PR 12-SEP-2002; 2002US-0410320P.  
 PR 16-SEP-2002; 2002US-0411060P.  
 PR 23-SEP-2002; 2002US-0412766P.  
 PR 24-SEP-2002; 2002US-0412825P.  
 PR 25-SEP-2002; 2002US-0413342P.  
 PR 30-SEP-2002; 2002US-0414832P.  
 XX (CURA-) CURAGEN CORP.  
 PA Zhong M, Guo X, Anderson DW, Ort T, Padigaru M, Rieger DK;  
 PI WPI; 2004-315567/29.  
 DR P-PSDB; ADL57159.  
 XX New isolated NOVX polypeptides and polymucleotides, useful for  
 PT preventing, diagnosing or treating NOVX-associated disorders, e.g.  
 PT osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,  
 PT asthma, or infections.  
 PS Claim 17; SEQ ID NO 103; 214pp; English.  
 CC The invention relates to a novel isolated polypeptide (NOVX) comprising a  
 CC mature form of any of the 37 amino acid sequences fully defined in the  
 CC specification. A polypeptide of the invention has antidiabetic,  
 CC anorectic, cardiac, hypotensive, antiatherosclerotic, anorectic,  
 CC virocidic, antibacterial, fungicide, protozoacidal, nootropic,  
 CC neuroprotective, antiparkinsonian, anticonvulsant, osteoprotective,  
 CC antiarthritic, antiinflammatory, dermatological, antisthmatic, and  
 CC antileptemic activity. A polymucleotide of the invention may have a use  
 CC in gene therapy. The polypeptides, nucleic acid molecules and antibodies  
 CC are useful in the manufacture of a medicament for treating a syndrome  
 CC associated with a human disease, preferably a NOVX-associated disorder.  
 CC The nucleic acid molecules, polypeptides and antibodies are useful for  
 CC treating, preventing or diagnosing diseases such as metabolic disorders,  
 CC diabetes, obesity, infectious diseases (viral, bacterial, fungal,  
 CC helminthic, and protozoal), anorexia, cancer, cardiovascular diseases  
 CC (hypercrenion, atherosclerosis), neurodegenerative disorders, Alzheimer's  
 CC disease, Parkinson's disease, epilepsy, immune disorders  
 CC (osteoarthritis), haematopoietic disorders, inflammatory skin disorders,  
 CC asthma, and various dyslipidaemias. The nucleic acids and polypeptides  
 CC may also be used as targets for the identification of small molecules  
 CC that modulate or inhibit e.g. neurogenesis, cell differentiation, cell  
 CC proliferation, haematopoiesis, wound healing and angiogenesis, in gene  
 CC therapy. In generation of antibodies that bind immunospecifically to NOVX  
 CC substances for use in therapeutic or diagnostic methods. The nucleic  
 CC acids are further used as hybridisation probes, in chromosome mapping,  
 CC tissue typing, preventive medicine, and pharmacogenomics. The NOVX  
 CC polypeptides of the invention show homology to certain known human  
 CC proteins: NOVXa-1c shows homology to fibroblast growth factor receptor 4  
 CC (FGFR4); NOVXa shows homology to complement factor I precursor; NOVXa  
 CC shows homology to matrix metalloproteinase-15 precursor; NOVXa shows  
 CC homology to MOC3; NOVXa-5c shows homology to T-lymphocyte surface antigen  
 CC LY-9 precursor; NOVXa-6m shows homology to fibroblast growth factor-21  
 CC (FGF-21); NOVXa-7c shows homology to alpha-2 macroglobulin-like  
 CC polypeptide variant; NOVXa-8g shows homology to antileukoprotease 1  
 CC precursor; NOVXa-9i shows homology to LIV-1 protein; NOVXa shows homology  
 CC to nuclear hormone receptor NOR-1; NOVXa-11j shows homology to  
 CC transmembrane protein-like; NOVXa-12c shows homology to beta-neoendorphin  
 CC dydorphin precursor. The present sequence encodes a NOVX polypeptide of  
 CC the invention.  
 XX  
 SQ Sequence 1101 BP, 283 A, 236 C, 262 G, 320 T, 0 U, 0 Other;  
 Alignment Scores:

Pred. No.:	2,31e-154	Length:	1101
Score:	1599.00	Matches:	302
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	12	Gaps:	0
US-10-659-004-104 (1-302) x ADL57158 (1-1101)			
QY 1	MethAlaArgLYLeuSerValIleLeuThrPheAlaLeuSerValThrAsnPro		20
DB 123	ATGCGAGAGAAATATCTGATCTTATCTGACCTTGGCCCTCTGTCACAAATCC		182
QY 21	LeuHISGLIeulYsAlaAlaPheProGlnThrThyGluYsIleSerProAsnTrp		40
DB 183	CTTCATGAACTMAAAGACGCTTCCCGACACCACTGAGAAATTTAGTCCAAATTGG		242
QY 41	GLuSerGlyIleAsnValaPheLeuAlaIleSerThrArgGlnTrpHisLeuGlnLeu		60
DB 243	GAATCTGGCATTAATGTTGACTTGGCAATTTCCACGGCAATTCATCTCAACAGCTT		302
QY 61	PheTYrArgTYrGlyGluAsnAsnSerLeuSerValGluGlyPheArgLYLeuLeuGln		80
DB 303	TTCTACCGCTATGAGAAATATTTCTTGTCACTTGAAGGTTCCAGAAATTAATTCTCA		362
QY 81	AsnIleGlyIleAspLYsIleLYsArgIleHisIleHisAspHisIleSer		100
DB 363	AATATAGCATAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA		422
QY 101	AspHisGluHisIleSerAspHisGluArgHisIleSerAspHisGluHisIleSerAspHis		120
DB 423	GACACAGACATCACTGACACATGACATGACATGACATGACATGACATGACATGACATG		482
QY 121	HisProHisSerHisSerGlnArgTYrSerArgGluGluLeuLYsAspAlaGlyValAla		140
DB 483	CATCTCAAGTCAACACGACGCTTCTCGGAGAGAGTGAAGATGCGGCGTGGCC		542
QY 141	ThrLeuAlaTrpMetValIleMetGlyAspGlyLeuHisAsnPheSerAspGlyLeuAla		160
DB 543	ACTTTGGCTGAGATGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG		602
QY 161	IleGlyAlaAlaPheThrGluGlyLeuSerSerGlyLeuSerThrSerValAlaValPhe		180
DB 603	ATTGCTGCTGCTTTTATCTGAAGCTTATCAAGTGTATTAAGTACTTCTGTTGCTGCTTC		662
QY 181	CysHisGluLeuProHisGluLeuGlyAspPheAlaValLeuLeuLYsAlaGlyMetThr		200
DB 663	TGTCATGATGCTTCTCTCATGAATTAAGTGAATTTCTGTTCTTAATTAAGTGAATG		722
QY 201	ValIleGlnAlaValIleLeuTYrAsnAlaLeuSerValMetLeuAlaTYrLeuGlyMetAla		220
DB 723	GTTAAGCAAGGCTGCTTCTTATTAAGCATTTGTCACCATGCTGCGATTTTGAATGCA		782
QY 221	ThrGlyIlePheIleGlyHisTYrAlaGluAsnValSerMetTrpIlePheAlaLeuThr		240
DB 783	ACAAGAAATTTTCAATGTCATTAATGCAAAATTTTCTAATGCAATTAATTTGCACTTACT		842
QY 241	AlaGlyLeuPheMetHisValAlaLeuValAspMetValProGluMetLeuHisAsnAsp		260
DB 843	GCTGCTTATTAATCAAGATGTTGCTGCTGATTAAGTGAATTAAGTGAATTAAGTGA		902
QY 261	AlaSerAspHisGlyCysSerArgTYrGlyTYrPhePheLeuGlnAsnAlaGlyMetLeu		280
DB 903	GCTATGATCACTGATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGA		962
QY 281	LeuGlyPheGlyIleMetLeuLeuIleSerIlePheGluHisGlyIleValPheArgIle		300
DB 963	TTGGGTTTGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA		1022
QY 301	AsnPhe 302		
DB 1023	AATTTT 1028		

RESULT 2	
ADNR28033	
ID	ADNR28033 standard; cDNA; 2250 BP.
XX	
AC	ADNR28033;
XX	
DT	04-NOV-2004. (first entry)
XX	
DE	Human LIV-1 protein encoding cDNA.
XX	
KW	Antibody; IGSF9; immunoglobulin superfamily member 9; LIV-1; neoplastic disorder; vaccine; anti-idiotypic; cancer; antisense therapy; cyostatic; gene; ss.
XX	
OS	Homo sapiens.
XX	
FT	Key
FT	Location/Qualifiers
FT	1..2250
FT	/tag=
FT	/product= "Human LIV-1 protein"
XX	
FN	MO200406933-A2.
XX	
PD	12-AUG-2004.
XX	
PF	27-JAN-2004; 2004WO-US002044.
XX	
PR	27-JAN-2003; 2003US-0442535P.
XX	
PA	(MCLA/) MCLACHLAN K.
PA	(GLAS/) GLASER S.
PA	(PEAC/) PEACH R J.
PA	(ROME/) ROME T.
PI	Mclachlan K, Glaser S, Peach RJ, Rowe T;
XX	
DR	WPI: 2004-580843/56.
XX	
DR	P-PsDB; ADNR28034.
XX	
PT	New antibody or its antigen binding fragment that associates with either IGSF9 or LIV-1 polypeptide, and the polypeptides, useful in preparing a vaccine composition for treating cancer.
PT	
XX	
PS	Disclosure; SEQ ID NO 28; 149bp; English.
XX	
CC	The invention relates to a novel isolated antibody or its antigen binding fragment that associates with either IGSF9 (immunoglobulin superfamily member 9) or LIV-1 polypeptide, and the fully defined polypeptides. The polypeptide comprises a sequence selected from ADNR28009, ADNR28011, ADNR28037, ADNR28028, ADNR28023, ADNR28030, ADNR28031, and ADNR28032, having 966 or 979 amino acids. The invention further comprises: a composition comprising the antibody which associates with IGSF9 or LIV-1 for treating a neoplastic disorder or the polypeptide comprising a sequence having 966 or 979 amino acids; a method of treating a mammal exhibiting a neoplastic disorder; a vaccine for treating cancer comprising the IGSF9 or LIV-1 polypeptide or an anti-idiotypic antibody that immunologically mimics the IGSF9 or LIV-1 antigen or its fragment and a carrier; a method of inducing an immune response in a patient in need of treatment or prevention of cancer; a method of diagnosing cancer by detecting overexpression of IGSF9 or LIV-1 or its fragment; a method for determining the prognosis of an individual receiving a cancer treatment; a kit comprising the composition together with instructions for use to treat or detect cancer; a method of treating a neoplastic disorder in a mammal where the neoplastic cells express the IGSF9 or LIV-1 antigens; an antisense nucleic acid up to 50 nucleotides in length comprising at least an 8 nucleotide portion of IGSF9 or LIV-1 which inhibits the expression of IGSF9 or LIV-1; a method of inhibiting the expression of IGSF9 or LIV-1 in cells or tissues; an isolated nucleic acid selected from the group consisting of: ADNR28008, ADNR28010, ADNR28017, ADNR28018, ADNR28019, ADNR28020, ADNR28021, ADNR28022, ADNR28023, ADNR28024, ADNR28025, and ADNR28026; a vector comprising the nucleic acid; and a host cell comprising the nucleic acid. The antibody has cytostatic activity. The antibody or its antigen binding fragment that associates with either IGSF9 or LIV-1

CC	polypeptide, and the IGF9 or IIV-1 polypeptides, are useful in preparing		
CC	a composition for treating cancer. This polynucleotide sequence		
CC	represents the cDNA encoding a human IIV-1 protein of the invention.		
XX			
XX	Sequence 2250 BP; 657 A; 503 C; 474 G; 616 T; 0 U; 0 Other;		
Alignment Scores:			
Pred. No.:	2,44e-129	Length:	2250
Score:	1359.50	Matches:	301
Percent Similarity:	40.32%	Conservative:	1
Best Local Similarity:	40.19%	Mismatches:	0
Query Match:	85.02%	Indels:	447
DB:	13	Gaps:	1
US-10-659-004-104 (1-302) x ADR28033 (1-2250)			
QY	1	MetAlaAqLyLeuSerValIleLeuIleLeuThrPheAlaLeuSerValThraenPro	20
Db	1	ATGGGAGGAAGACTTATCTGTATCTTGATCCGACCTTGCCCTCTCTGCACAAATCCC	60
QY	21	LeuHisGluLeuLysAlaAlaAlaPheProGlnThrGluLysIleSerProaenTP	40
Db	61	CTTCATGAACATPAAAGCAGAGCTGTTCCCCAGACACAGAAATATGTCGAATTGG	120
QY	41	GIuSerGylIleAsnValAspLeuAlaIleSerThrArgGluThyHisLeuGlnIleu	60
Db	121	GAATCTGGCATTTAAAGTGGACTTGGCAATTTCCACCGGCAATATCATCTACAAACGCTT	180
QY	61	PheTYArgTYGlyGluLysAsnAsnSerLeuSerValGluGlyPheArgLysLeuLeuGln	80
Db	181	TTCTACCGCTATGGAGAAATATATCTTTGTGACTTGAAGGGTTCAGAAATATTACTCA	240
QY	81	AsnIleGlyIleAspLysIleLysArgIleHisIleHisHisAspHisAspHisSer	100
Db	241	AATATAGGATGATGATTAAGATTAAAGAAATCCATATACACATGACACGACCATCACTCA	300
QY	101	AspHisGluHisHisSerAspHisGluAlaArgHisSerAspHisGluHisHisSerAsp	119
Db	301	GAACGACGAGCATCACTCAGACCATGAGCGTCACTAGACCATGAGCATCATCAGACAC	360
QY	119	-----	119
Db	361	GAGCATCACTTGACCATATATCATGCTGTTGTGTAATAATAGCGAAAGCTCTTTC	420
QY	119	-----	119
Db	421	CCAGACATGACTCATGATAGTTCAAGTAAGATCTTGAAGACGACGAGGAAAGAGCT	480
QY	119	-----	119
Db	481	CACGACGAGAACATGCGCAGTGTAGAGAAATGTCAAGACAGATGTTAGTCAAGTAA	540
QY	119	-----	119
Db	541	GTGACCTCAACTGTGTACAACACTGTTCTGTGAAGAACTCACTTTTGAAGACAATAGAG	600
QY	119	-----	119
Db	601	ACTCGAAGCCTGGAATACTTTCGCCAAGAATGTATAGACAGCTCACTCCACCCAGTGC	660
QY	119	-----	119
Db	661	ACATCAAGAGCCGGGTGAGCCGGCTGCTGTGAGAAAACAATGAAATCTGTGAGTGAG	720
QY	119	-----	119
Db	721	CCCGAAAAAGCTTATGTATTTCCAGAAAAACAATGAAATTCCTCAGAGGTTCAT	780
QY	119	-----	119
Db	781	GCATCAAGCTACTGACATCTCATGGCATGGGCGATCCAGTTCCGCTGAATGCAACAGAG	840
QY	119	-----	119

```
Db      841 TTCAACTATCTCTGTCCAGCCATCATCAACCAATTGATGCTTGTCTGATTCAT 900
Qy      119 -----
Db      901 ACAAGTGAAGAAGGCTGAATCCCTCCAAAGACCTATTCATTACAATAGCCGGTT 960
Qy      119 -----
Db      961 GGTGGTTTATAGCCATTTCATCATCAGTTTCCTGCTGCGGGGTTATCTTAGTG 1020
Qy      119 -----
Db      1021 CCTCTCATGAATCGGGGTGTTTCAAAATTTCTCTGAGTTTCCTTGCGCACTGGCCGTT 1080
Qy      119 -----
Db      1081 GGGACTTTGAGTGTGATGCTTTTTCACACTTTCACATTCATCATGCAAGTCAGCAC 1140
Qy      119 -----
Db      1141 CATAGTCATAGCCATGAAAGAACAGCAATGAAATGAAAGAGACCACTTTTCAGTCAT 1200
Qy      119 -----
Db      1201 CTGTCTTCTCAAAACATAGAAAGAAAGTGCTTATTTGATTCACAGTGGAGGTTCAACA 1260
Qy      119 -----
Db      1261 GCTTAGAGAGCCGTGATTTTCATGTTCTTGTGAACATGTCCTCATTCATTCAAACAA 1320
Qy      119 -----
Db      1321 TTTAAAGATTAAGAGAAAAAGATCAAGAAAGAACTGAAATGATGATGTGGAGATT 1380
Qy      119 -----
Db      1381 AAGAGCAGTTGTCCAAGTATGATCTCAACTTTCACAAATGAGAGAAAGTAGATACA 1440
Qy      119 -----
Db      1441 GATGATCGAAGTGAAGGCTATTACGAGCAGACTCAAGAGCCCTCCCATTTGATTCAT 1500
Qy      119 -----
Db      1501 CAGCAGCCTGCAGTCTTGAGAGAAAGAGTCAATGATCTCATCTCAGCAGAGAA 1560
Qy      119 -----
Db      1561 GTCTACATGATATGTACCCAGAGGTCGCAAGATTAATGCCATTTCACATTTCCAGAT 1620
Qy      119 -----
Db      1621 ACACCTGGCCAGTCAGACGATCTCATTCACCAACATCATGACTACATCATATTCCTCAT 1680
Qy      120 -----
Db      1681 CATCACCAACCAAAACCAACCATCTCTCAGCTCAAGCCAGCTACTCTCGGAGAGAG 1740
Qy      134 Leu1yAspAlaGlyValAlaThrLeuAlaTrpMetVal11leMetGlyAspGlyLeuHis 153
Db      1741 CTGAAGAAGATGCCGGGTGCGCACTTTGGCTGGATGATATAGGATGAGTGGCTGCAC 1800
Qy      154 AsnProSerAspGlyLeuAla11leGlyValAla1aPheThrGluGlyLeuSerSerGlyLeu 173
Db      1801 AATTTCAGGATGGCTTACGCAATGTGTGCTTCTTACCTGAAGGCTTATCAAGTGGTTA 1860
Qy      174 SerThrSerValAlaValaPheCysHisGlyLeuLeuProHisGlyLeuGlyAspPheAlaVal 193
Db      1861 AGTACTTCTGTTGCTGTGCTGTCTGATGAGTGGCTTCATGAATTCGGTGAATTCCTGTT 1920
Qy      194 LeuLeuValaGlyMetThrVala1ySglnAlaVala1leuTyAsnAlaLeuSerAlaMet 213
```

```
Db      1921 CTACTAAAGGCTGGCAGTACCGCTTAAGCAGAGCTGCTCCTTATATATGATTCAGCCATG 1980
Qy      214 LeuAlaTyLeuGlyMetAlaTrpGly11lePhe11leGlyHisTyAlaGlyAsnValSer 233
Db      1981 CTGGCGTATCTTGGAAATGGCAACAGAAATTTTCATGTGATATTCATGCAAAATGTTTCT 2040
Qy      234 MetTrp11lePheAlaLeuThrAlaGlyLeuPheMetHisVala1AlaLeuValaAspMetVal 253
Db      2041 ATGTGATATTTGCACTTACTGCTGGCTTATCTCATGATATGCTCTGTGTGATATGATA 2100
Qy      254 ProGluMetLeuHisAsnAspAlaSerAspHisGlyCysSerArgTrpGlyTyPhePhe 273
Db      2101 CTGAAATGCTGCACATGATGATGATGACCATGATGATGACCGCTGGGGTATTTCTTT 2160
Qy      274 LeuGlnAsnAlaGlyMetLeuLeuGlyPheGly11leMetLeuLeu11Ser11lePheGlu 293
Db      2161 TTACAGATGCTGGGAGTCTTTGGGTTTGAATATGTTACTTATTCATATTTGAA 2220
Qy      294 HisTyAla11leValaPheArg11leAsnPhe 302
Db      2221 CATAAATCGTGTTCGTATTAATTTTC 2247

RESULT 3
ACC50210 standard; cDNA; 2744 BP.
ACC50210;
12-JUN-2003 (first entry)
DE Breast cancer associated cDNA sequence SEQ ID NO:263.
XX Human; breast cancer; cytostatic; gene therapy; gene; ss.
XX Homo sapiens.
XX MO2003004989-A2.
XX 16-JAN-2003.
XX 21-JUN-2002; 2002MO-US019669.
XX 21-JUN-2001; 2001US-0299887P.
XX 27-JUN-2001; 2001US-0301572P.
XX 18-JUL-2001; 2001US-0306501P.
XX 25-SEP-2001; 2001US-0325002P.
XX 05-MAR-2002; 2002US-0362585P.
XX 14-MAY-2002; 2002US-0380391P.
XX (MILL-) MILLENIUM PHARM INC.
XX Lillie J, Gannavarapu M, Glatt K, Hoersh S, Kamackar S,
XX Mertens M, Monahan JE, Myer V, Wang Y, Xu Y, Zhao X, Meyers RE;
XX Baat RC, Hortobagyi GN, Fuzsai L, Meric F, Sahin A, Mills GB;
XX WPI; 2003-210381/20.
XX P-PSDB; ABR47514.
XX Breast cancer diagnosis or treatment by comparing the level of expression
XX of a marker in a patient sample with that in the control non-breast
XX cancer sample.
XX Claim 1; SEQ ID NO 263; 128bp; English.
XX The present invention describes a method for assessing whether a patient
XX is afflicted with breast cancer. The method comprises comparing the level
XX of expression of a marker (gene/polypeptide see ACC50076 to ACC50334 and
XX ABR47286 to ABR47632) in a patient sample and the normal level of
XX expression of the marker in a control non-breast cancer sample, where a
XX significant increase in the level of expression of the marker in the
XX patient sample and the normal level is an indication that the patient is
XX afflicted with breast cancer. The breast cancer associated sequences from
XX the present invention have cytostatic activities and can be used in gene
```

CC therapy. The method is useful for diagnosing and treating breast cancer.  
CC N.B. The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pat\_sequences

XX Sequence 2744 BP; 786 A; 585 C; 601 G; 772 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.:	3,19e-129	Length:	2744
Scores:	1359.50	Matches:	301
Percent Similarity:	40.32%	Conservative:	1
Best Local Similarity:	40.19%	Mismatches:	0
Query Match:	85.02%	Indels:	447
DB:	8	Gaps:	1

US-10-659-004-104 (1-302) x ACCS0210 (1-2744)

```
QY 1 MetAlaArgLyLeuSerValIleuLeuThrPheAlaLeuSerValThrAsnPro 20
Db 138 ATGGCGAGAGATTATCTGTAACTTGAATCCTGACCTTGGCCCTCTCTGTACAAATCCC 197
QY 21 LeuHiegluLeuValAlaAlaPheProGlnThrGluValIleSerProAsnTrp 40
Db 198 CTTGATGAATAAAGCAGCTGCTTCCCGACGACCTAGAGAAATTAATGCGAATTGG 257
QY 41 GluSerGlyIleAsnValAspLeuAlaIleSerThrArgGlnTrpHisLeuGlnLeu 60
Db 258 GAATCTGGCATTAATGTTGACTTGGCAATTTCCACGCGCAATTCATCTCAACAGCTT 317
QY 61 PheTyArgTyArgLygluAsnAsnSerLeuSerValGluGlyPheArgLyLeuLeuGln 80
Db 318 TTCTACCCGCTATGAGAAATTAATTTCTTGTCACTGAGGGTTCAAGAAATTAATCTTCA 377
QY 81 AsnIleGlyIleAspLyLeuValIleValArgIleHisIleHisIleAspHisIleSer 100
Db 378 AATATAGGCATAGATTAAGATTAAAGAAATCCATATACCATGACGACGACATCATCTCA 437
QY 101 AspHisGluHisIleSerAspHisIleGluArgHisIleSerAspHisIleSerAsp 119
Db 438 GACCACGAGCATCACTGACCATGAGCGTCACTCAGACCATGAGCATCACTCAGACCCAC 497
QY 119 ----- 119
Db 498 GAGCATCACTGACCATATCATATGCTGCTTCTGTAAATAAAGCGAAAGCTTTTGC 557
QY 119 ----- 119
Db 558 CCAAGACATGACTGAGATAGTTCAAGTAAAGATCTCTAAGAACGCCAGGGGAAGAGCT 617
QY 119 ----- 119
Db 618 CACGACGAGAACATGCGAGTGTAGAAGATGTCAGAGACAGTGTATGCTAGTCAA 677
QY 119 ----- 119
Db 678 GTGACTTCACTGTGTACACACTGTCTTGAAGAACTCACTTTTAGAGACAATAGAG 737
QY 119 ----- 119
Db 738 ACTCCAAAGACTGGAAAACTTCCCAAAAGATGTAAAGACAGCTCCACCCAGTGTCT 797
QY 119 ----- 119
Db 798 ACATCAAGAAGCCGGGTGAGCCGGCTGGTGTAGAAACAAATGAATCTGTGAGTGA 857
QY 119 ----- 119
Db 858 CCCCCAAAAGGCTTTATGTATTCAGAAACACAAATGAATTCCTCAGAGAGTGTTCAT 917
QY 119 ----- 119
Db 918 GCATCAAGTACTGACATCTCATGGCATGGGATCCAGTTCCGCTGAATGCAACAGAG 977
```

```
QY 119 ----- 119
Db 978 TTCACTATCTGTGTCCAGCCATCATCAACCAATGTAGTCTGTGATTCAT 1037
QY 119 ----- 119
Db 1038 ACAAGTGAAGAAGGCTGAATCCCTCCAAGACCTATTCAATTAAGCTGGGTT 1097
QY 119 ----- 119
Db 1098 GGTGTTTATATAGCATTTCCATTCATCAGTTTCTGTCTGTGCGGGTATCTTAGTG 1157
QY 119 ----- 119
Db 1158 CTTCTCATGAATGGGGTGTGTTTTCAAATTTCTCTGAGTTTCTTGTGGCACTGGCGTT 1217
QY 119 ----- 119
Db 1218 GGGACTTGAAGTGTGATGCTTTTACACCTTCTTCCATTCATGCAAGTCACAC 1277
QY 119 ----- 119
Db 1278 CATAGTCATAGCATGAAGAACAGCAATGGAATGAAGAAAGAGACACTTTTCAGTCAT 1337
QY 119 ----- 119
Db 1338 CTGCTTCTCAAAACATAGAAAGAAAGTCTTATTTGATTCACGTGAAGGGTCTAACA 1397
QY 119 ----- 119
Db 1398 GCTCTAGAGGCGCTGATTTTCATGTTCTTGTGAAATGTCCTCAGATTGATGAACAA 1457
QY 119 ----- 119
Db 1458 TTTAAGATGAAGAAAGAAAGAAATCAGAGAAACCTGAAATGATGATGTGAGATT 1517
QY 119 ----- 119
Db 1518 AAGAGAGATTGCCAAGTATGAATCTCACTTTCAACAAATGAGAGAAAGTAGATACA 1577
QY 119 ----- 119
Db 1578 GATGATGAAGTGAAGGCTATTTACGAGACACTCAAGAGCCCTCCACTTTGATTC 1637
QY 119 ----- 119
Db 1638 CAGAGCCTGCAATCTTGAAAGAAAGAGGTATGATAGTCACTGCTACACAGAA 1697
QY 119 ----- 119
Db 1698 GTCTACATGAATATGTATCCAGAGGCTGCAGAAATGAATGCCATTCATTTCCAGAT 1757
QY 119 ----- 119
Db 1758 ACACTGGCCAGTCAGACATCTCATCCACCATATGACTACATCATTTCTCCAT 1817
QY 120 ----- 119
Db 1818 CATCACACACCAAAACCAACCATCTCACAGTACAGCCAGGCTACTCTCGGAGAGAG 1877
QY 134 LeuLyAspAlaGlyValAlaThrLeuAlaTrpMetValIleMetGlyAspGlyLeuHis 153
Db 1878 CTGAAGATCCCGCGCTGCGCACCTTGGCTGTGATTAAGGGTGTATGGCTGCAC 1937
QY 154 AsnPheSerAspGlyLeuAlaIleGlyAlaAlaPheThrGluGlyLeuSerSerGlyLeu 173
Db 1938 AATTTCCGATGGCTTAGCAATTTGGTGTCTCTTTTACTAAGGCTTATCAAGGCTTTA 1997
QY 174 SerThrSerValAlaValPheCyHisGlyLeuProHisGluLeuGlyAspPheAlaVal 193
Db 1998 AGTACTTCTGTGTGTCTGTCTGTATGAGATGGCTCATGAATTAAGTGACTTGTGCTT 2057
QY 194 LeuLeuValaGlyMetThrValLyGlnAlaValLeuTrpAsnAlaLeuSerAlaMet 213
```

Db 2058 CTTACTAAAGCGTCGATGACCGGTTTAGCAGCGCTGCTTTTAAATGACATTGTCAGCCATG 211

Qy 214 LeuAlaTYLeuGIyMeAlaThrGIyIlePheIIeGIyHsTYAlGIuAsnValSer 233

Db 2118 CTGGCGATCTCTGGAAATGACCAACAGGAATTTTCATTGGCATTTATGCTGAAATGTTTCT 217

Qy 234 MetTrpIlePheAlaLeuThrAlaGIyLeuPheMetHsValAlaLeuValAspMetVal 253

Db 2178 ATGTGGATATTGTGCACTACTGCTGCTTATTCATGATGATGTTGCTTGGTGGATATGGTA 223

Qy 254 ProGUwMeLeuHsAsnAspAlaSerAspHsGIyCySserArgTrpGIyTYrPhePhe 273

Db 2238 CTTGAAATGCTGCACATGATGCTTGTACCACTGATGATGAGCCGCTGGGGGATTTCTT 229

Qy 274 LeuGlnAsnAlaGIyMeLeuLeuGIyPheGIyIleMeLeuLeuIleSerIlePheGlu 293

Db 2298 TTACAGATGCTGGGATGCTTTGGGTTTGGAAATTATGTACTTATTTCCATATTGGAA 235

Qy 294 HisIysIleValPheArgIleAsnPhe 302

Db 2358 CATAAATCGTGTTCGTATTAATTTTC 2384

RESULT 4

ADH28902

ID ADH28902 standard; DNA; 2744 BP.

XX ADH28902;

AC

XX 11-MAR-2004 (first entry)

DT

XX Human chronic myelogenous leukemia (CML) gene marker #170.

DE

XX ds; chronic phase chronic myelogenous leukaemia; CP-CML;

KW blast crisis CML; BC-CML; human; chronic myelogenous leukaemia;

KW gene marker.

XX

XX Homo sapiens.

OS

PN US2003104426-A1.

XX

PD 05-JUN-2003.

PF 14-JUN-2002; 2002US-00171581.

XX

PR 18-JUN-2001; 2001US-0298914P.

XX

PA (LINS/) LINSLEY P S.

XX

PA (MAOM/) MAO M.

PA (DATH/) DAI H.

PA (HEY/) HE Y.

PA (RADI/) RADICH J P.

XX

PI Linsley PS, Mao M, Dai H, He Y, Radich JP;

XX

DR WPI; 2003-787046/74.

XX

PT Classifying cell sample as chronic phase chronic myelogenous leukemia or

PT blast crisis chronic myelogenous leukemia by detecting difference in

PT expression of genes corresponding to the markers such as X15415, U89436.

XX

PS Disclosure; SEQ ID NO 170; 31pp; English.

XX

CC The invention relates to a method of classifying a cell sample as chronic

CC phase chronic myelogenous leukaemia (CP-CML) or blast crisis CML (BC-

CC CML). The method is useful for classifying a sample as CP-CML or BC-CML.

CC The present sequence represents a human chronic myelogenous leukemia

CC (CML) gene marker used to distinguish blast crisis CML from chronic phase

CC CML.

XX

XX Sequence 2744 BP; 786 A; 585 C; 601 G; 772 T; 0 U; 0 Other;

XX

Alignment Scores:

Record No.:	3,19e-129	Length:	2744
Score:	1359.50	Matches:	301
Percent Similarity:	40.32%	Conservative:	1
Best Local Similarity:	40.19%	Mismatches:	0
Query Match:	85.02%	Indels:	447
DB:	10	Gaps:	1
US-10-659-004-104 (1-302) x ADH28902 (1-2744)			
QY	1 MetalArglyLeuSerValIleLeuIleLeuTherPhealLeuSerValThrAspPro	20	
Db	138 ATGCGCAGGAAGTTATCTGTAACTTGATCTTGATCCGACCTTGCCTCTGTGCACAAATCCC	197	
QY	21 LeuHisgluLeuysalaaalaphProGlnThrThrgluysIleSerProAsnTrp	40	
Db	198 CTTGATGAACTAAAGCAGCTGCTTTCGCCAGACCACTGAAATTAATGTCGAAATGG	257	
QY	41 GluSerGlyIleAsnValAspLeuValIleSerThrArgInTrnIleuGlnGlnLeu	60	
Db	258 GAATCTGGCATTAATGTGACTTGCAATTTCCACAGGCATATCATCTCAACAGCTT	317	
QY	61 PheTrpArgTrgIlyGluAsnAsnSerLeuSerValGluGlyPheArglyLeuLeuGln	80	
Db	318 TTCTACCGCTATGGAGAAATAATCTTGTCTGAGTTGAAGGCTTCAGAAATTAATCTTCA	377	
QY	81 AsnIleGlyIleAspIlyseIlyAsxGlyIleHisIleHisAspHisAspHisSer	100	
Db	378 AATATAGGCATAGTAATGATTAAGAATTCATATACCATCCATGACACGACCATCATCTCA	437	
QY	101 AspHisgluHisHisSerAspHisgluArgHisSerAspHisgluHisHisSerAsp	119	
Db	438 GACCAACAGCATCTCTCAGACACATGACGCTACTCAGACCATGAGCATCACTACAGCAC	497	
QY	119 -----	119	
Db	498 GAGCATCATCTGACCAATATCATGCTGCTTCTGTAAATAAGCGAAAGCTCTTGTGC	557	
QY	119 -----	119	
Db	558 CCAGACCATGACTCAGATAGTTTCAGGTAAGATCTTGAAACAGCCAGCGGAAAGAGCT	617	
QY	119 -----	119	
Db	618 CACCGACCAGAAACATGCCAGTGAGTAAGAATTCACAGACAGCTAGTGTAGTGAA	677	
QY	119 -----	119	
Db	678 GTGACATCACTGTGTACAAACACTGTCTCTGAGAACTCACTTTCTAGAGCAATAGAG	737	
QY	119 -----	119	
Db	738 ACTCCAAAGCCTGGAATACTTCTCCCAAGATGTAAGCAGCTCCATCCACCCAGTGTG	797	
QY	119 -----	119	
Db	798 ACATCAAAAGACCCGGGTGACGGGCTGTGTGAGAAACAATAATGAATCTGTGAGTGAG	857	
QY	119 -----	119	
Db	858 CCCGAAAGGCTTATGTATATTCAGAAACAAATGAAAAATCTTCAGAGAGTTTCAAT	917	
QY	119 -----	119	
Db	918 GCATCAAAAGTACTGACATCTCATGCGATGGGCATCCAGGTTCCGCTGAATGCAACAGAG	977	
QY	119 -----	119	
Db	978 TTCAACTATCTCTGTCCAGCATCATCAACCAATTAATGAATCTTGTGTATTCAT	103	
QY	119 -----	119	
Db	1038 ACAAGTGAAGAAAGGCTGAATCCCTTCGAAAGCACTATTGATTAACAATATGCTGGGTT	1097	



QY	21	LeuHISGluLeuLeuYAlAAlaAlaPheProGlnThrThnGluLysAlIeserProAntTP	40
Db	198	CTTCAATGAACATAAAGCAGCTGCTTCCCCAGACACATGAGAAATTAAGTCCGAATTGG	257
QY	41	GIuSerGlyIlaAsnValaPheLeuAlaIleSerThrArgGlnTrpHisLeuGlnGlnLeu	60
Db	258	GAATCTGGCATTAATGTTTAACTTTGGCAATTTCCACCGGCAATATATCTTACAACAGCTT	317
QY	61	PheTyArgTyrgLyGluAsnAsnSerLeuSerValGluGlyPheArgLysLeuLeuGln	80
Db	318	TTCTACCCGATAGAGAAATATATCTTTGTGACTGGAAAGGGTTACGAAATTAATTACTTCAA	377
QY	81	AsnIleGlyIlaAspLysIleLysValGllIleHisIleHisIleAspHisAspHisIleSer	100
Db	378	AATATAGCGATGATTAAGTAAATTAAGAAATCCATATACACCATACCAACGACCATCTCA	437
QY	101	AspHisGluHisIleSerAspHisGluValArgHisIleSerAspHisGluHisIleSerAsp	119
Db	438	GACCAAGACATCACTACAGACCATGAGCGTCACTCAGACCATGATGATCATCTGACACAC	497
QY	119	-----	119
Db	498	GAGCATCACTGACCATATATCATGTGCTTCTGGTAAATAAGGAAAGCTCTTTGC	557
QY	119	-----	119
Db	558	CCAGACCATGACTCAGATAGTTACAGTTAAAGATCTTGAACACGACGAGGAAAGAGCT	617
QY	119	-----	119
Db	618	CACCGACGAAACATGCCAGTGGTAGAAGAAATGTCAAGACAGTGTTAAGTCTAGTAA	677
QY	119	-----	119
Db	678	GTGACCTCACTGTGTACAACACTGTCTGAAAGAACTCATCTTCTAGAGACAATAGAG	737
QY	119	-----	119
Db	738	ACTCAAGAACCTGAAAACTTCCCCAAAGATGTAAGAGCTCCACTCCACCCAGTGC	797
QY	119	-----	119
Db	798	ACATCAAGAGCCGGGTGAGCCGGCTGGCTGTGAGAAAACAATGATCTGTAGTGA	857
QY	119	-----	119
Db	858	CCCCGAAAGGCTTATATGATTTCCAAACACAAATGAAAAATCTCAGAGTGTTCAT	917
QY	119	-----	119
Db	918	GCATCAAAAGCTACTGACATCTCATGGCATGGGATCCAGTTCGGCTGAATGCAACAG	977
QY	119	-----	119
Db	978	TTCAACTATCTGTCCAGCCATCATCAACCAATTTGATGTAGATCTTGTCTGATTCAT	1037
QY	119	-----	119
Db	1038	ACAAAGTAAAGAGGCTGAATCCCTCCAAAGACTATTCATTACAAATAGCTGGTT	1097
QY	119	-----	119
Db	1098	GGTGGTTTATAGCATTTTCATCATCAGTTTCTGTCTCTGCTGGGGGTTATCTTAGTG	1157
QY	119	-----	119
Db	1158	CCTCTCATGAATCGGGTGTTTTTCAAATTTCTCTGAGTTTCTTGTGGCATGGCCGTT	1217
QY	119	-----	119
Db	1218	GGGACTTTGATGTGATGCTTTTTTTACACCTTCTTCCACATTTCTATGCAAGTCAACC	1277
QY	119	-----	119

Db	1278	CATGTCATAGCCATGAAAGAA	CCAGCAATGAGAA	TGAAAAGAGAACCACTTTTCAGTCAT	133
QY	119	-----	-----	-----	119
Db	1338	CTGTCTTCTCAAAA	CATAGAGAAAGTCCAT	TTTGTATTCACGCTGAGAGGCTTACA	1397
QY	119	-----	-----	-----	119
Db	1398	GCTCTAGAGAGCCTGTAT	TTTCATGTTTCTTGTTGAA	CATGTCCTCATTTGATCAACAA	1457
QY	119	-----	-----	-----	119
Db	1458	TTTAAAGATAGAAAGAAAGAA	GATCAGAGAAACCTGAAA	TGATGATGATGAGATT	1517
QY	119	-----	-----	-----	119
Db	1518	AAGAAAGAGTTGTCCAG	TATGATGATTCACCTTTCAC	CAATGAGAGAAAGTAGATACA	1577
QY	119	-----	-----	-----	119
Db	1578	GATGATCGAAGCTAGAGGCT	ATTATTCAGAGAGATC	CAGAGCCCTCCACTTTGATTCT	1637
QY	119	-----	-----	-----	119
Db	1638	CAGCAGCCTGCAGTCTTG	AGAAAGAAAGAGTCA	TGATAGCTCATGCTCATCCACAGAA	1697
QY	119	-----	-----	-----	119
Db	1698	GTTTACATGAATATGTAT	CCAGAGAGGTGCAAGAT	AAATGCAATTCACATTTCCACGAT	1757
QY	119	-----	-----	-----	119
Db	1758	ACACTGGCCAGTCAGAGAT	CTCATTCACCATCATGACT	CAATCATATTTCTCCAT	1817
QY	120	-----	-----	-----	133
Db	1818	CATCACCAACCAAAAC	CACCATCTCTCACAGT	CACAGCCAGGCTACTCTCGGAGAGAG	1877
QY	134	LeuLybAaPaIaGIyValaIa	ATThLeuAaIaTTPMeTValaI	ImeMeGIyAaPbGIyLeuHIs	153
Db	1878	CTGAAAGATGCCGGCGTGC	CCACTTGGCCCTGGAATGG	GATTAATGGGATGAGCCCTGCAC	1937
QY	154	AaPheSeAaPbGIyLeuAaI	IaLeGIyAaIaIaPheThrGIu	GIyLeuSeAaPbSeGIyLeu	173
Db	1938	AATTTCAAGCCTAGGCTT	AGCAATGGTGTCTTTACTGA	AGGCTTTATCAAGTGGTTTA	1997
QY	174	SeThrSeValaIaValaI	PheCyAHisGIuLeuProHIs	GIuLeuGIyAaPbPheAaIaI	193
Db	1998	AGTACTTCTGTGCTGTGT	CTGTATGATGTTGCTCATGA	ATTAGGTGACCTTGTGCTGT	2057
QY	194	LeuLeuLybAaIaGIyMe	ThValaIyGIyGlnAaIaVal	LeuLyTzaAaIaIaLeuSeAaIaI	213
Db	2058	CTACTAAAGCGTGCATGA	CCGTTTAAGAGGCTGTCT	TTTAATGCAATGTCCAGCCATG	2117
QY	214	LeuAaIyTLeuGIyMeAaI	aThThGIyIlePheIleGIyHIs	TrYzAaIaGIyAaAaValaISe	233
Db	2118	CTGGCGTATCTTGGAAT	GGCAACGAAATTTTCACTT	GGTCACTTATGCTGAAAATCTTTCT	2177
QY	234	MeTTrpIlePheAaIaLeu	ThAaIaGIyLeuPheMeThIs	ValaIaIaLeuValaIaPbMeValaI	253
Db	2178	ATGTGATATATTGCACT	TACTGTGCGCTTATTCATG	ATGTATGTTGCTGTGCTGTGATATG	2237
QY	254	ProGIyMeTLeuHIsAaA	aPbAaIaISeAaPbHIsGIyCy	SeSeAaATGTTrpGIyTTrPhe	273
Db	2238	CCTGAAAAGCTGCACAA	ATGATGATCAAGTACCAT	GGATGTATACCCGTGGGGATTTCTTT	2297
QY	274	LeuGIaAaAaIaGIyMe	LeuLeuGIyPheGIyIleMe	TLeuLeuLeuLeuLeuLeuLeuLeuLeu	293
Db	2238	TTACAGATGCTGGAGAT	CTTTTGGGTTTGGAATT	ATTAATTAATTTTCCATATTGCA	235
QY	294	HisLybIleValaI	PheATrGIleAaPbHe	302	



```

QY 119 ----- 119
Db 1218 GGGACTTGTAGTGTGAGCTTTTACACCTTCTTCCATCTTCATGCAAGTCAACAC 1277
QY 119 ----- 119
Db 1278 CATAGTCATAGCCATGAGAACACAGCAATGAAATGAAAAGAGACCACTTTTCAGTCAT 1337
QY 119 ----- 119
Db 1338 CTGTCTTCTCAAAACATGAGAGAAAGTGCCTATTGATTCCACGTGAGAGGGTCTAAC 1397
QY 119 ----- 119
Db 1398 GGTCTAGAGAGCGCTGATTTCATGTTTCTTGTGAACATGCTCCTCATGATGATCAACAA 1457
QY 119 ----- 119
Db 1458 TTTAAAGATAGAGAGAAAGAAATGAGAAAGCTGAAATGATGATGATGAGAGATT 1517
QY 119 ----- 119
Db 1518 AAGAAAGCACTGTCCAGATGATGATCTCACTTCAACAAATGAGAGAGAAAGTATACA 1577
QY 119 ----- 119
Db 1578 GATGATCGAAGCTGAAAGCTATTTCAGAGACACTCAGAGAGCCCTCCCATTTGATTCT 1637
QY 119 ----- 119
Db 1638 CAGAGCCTGAGTCTTGAGAGAGAGAGTATGATGATGATGATGATGATGATGATGATG 1697
QY 119 ----- 119
Db 1698 GTCTAATGAAATATGTACCCAGAGGTGCAAGAAATGATGATGATGATGATGATGATGAT 1757
QY 119 ----- 119
Db 1758 ACACCTCGGCACTAGAGACATCTCATTCACCATCATGATGATGATGATGATGATGATGAT 1817
QY 120 ----- 119
Db 1818 CATCAACCAACCAAAACCAACCATCTCTCAGCTCAGAGCAGCGCTACTCTGAGAGAGAG 1877
QY 134 LeuLYsAAspAlaGlyValAlaThrLeuAlaTrpMetValIleMetGlyAspGlyLeuHis 153
Db 1878 CTGAAAGATGCGCGGCGTGCCTTGGCTGATGATGATGATGATGATGATGATGATGATGATG 1937
QY 154 AamPheSerAspGlyLeuAlaIleGlyAlaAlaPheThrGluGlyLeuSerSerGlyLeu 173
Db 1938 AATTCAGAGATGCGCTAGCAATGTGTGCTTTTACTGAAAGCTTATCAAGTGTGTTA 1997
QY 174 SerThrSerValAlaValAlaPheCysHisGlyLeuPheProHisGlyLeuGlyAspPheAlaVal 193
Db 1998 AGTACTTCTGTGTGCTGTGTCTGTATGATGATGATGATGATGATGATGATGATGATGATG 2057
QY 194 LeuLeuLYsAAspAlaGlyMetThrValLySGlnAlaValLeuLYsAspAlaLeuSerAlaMet 213
Db 2058 CTACTAAAGAGCTGGCATGACCGCTTAAAGAGCTGCTCTTATTAATGCAATGTCACACCA 2117
QY 214 LeuAlaLYsLeuGlyMetAlaThrGlyIlePheIleGlyHisLYsLYsAlaGlyAsnValSer 233
Db 2118 CTGGCGTATCTTGGAATGCGCAACAGGAATTTTCATTTGGCTATATGCTGAAATGTTTTCT 2177
QY 234 MetTrpIlePheAlaLeuThrAlaGlyLeuPheMetHisValAlaLeuValAspMetVal 253
Db 2178 ATGTGGAATATTGCACTTACTGCTGCTTATTCATGATGATGATGATGATGATGATGATGAT 2237
QY 254 ProGluMetLeuHisAsnAspAlaSerAspHisGlyCysSerSerArgTrpGlyLYsPhePhe 273
Db 2238 CCGTAAATGCGTCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2297
QY 274 LeuGlnAsnAlaGlyMetLeuLeuGlyPheGlyIleMetLeuLeuIleSerIlePheGlu 293

```

```

Db 2238 TTAACGAAATGCTGGAGATGCTTTTGGGTTTGGATTATGTTACTTATTTCCATATTGAA 2357
QY 294 HisLYsIleValIlePheArgIleAsnPhe 302
Db 2358 CATAAATCGTGTTCGTATATAATTTC 2384

RESULT 7
ID ADR25539 standard; DNA; 2744 BP.
AC ADR25539;
DT 21-OCT-2004 (first entry)
DE Breast cancer prognosis marker #1400.
ds; breast cancer; prognosis; gene expression; diagnosis.
OS Homo sapiens.
PN W02004065545-A2.
PD 05-AUG-2004.
PE 15-JAN-2004; 2004WO-US001100.
PR 15-JAN-2003; 2003US-00342887.
XX (ROSE-) ROSETTA INPHARMATICS LLC.
XX (NECA-) NETHERLANDS CANCER INST.
XX Van't Veer LJ, He Y;
XX WPI; 2004-593473/57.
XX
XX Classifying a breast cancer patient according to prognosis comprises
XX determining the similarity between the level of expression of each of
XX five genes in a cell sample taken from patient, to control levels.
XX
XX Disclosure; SEQ ID NO 1400; 226bp; English.
XX
XX The invention relates to a method of classifying a breast cancer patient
XX according to prognosis by determining the similarity between the level of
XX expression of each of five genes for which markers are listed in the
XX specification, in a cell sample taken from the breast cancer patient, to
XX control levels of expression for each respective five genes to obtain a
XX patient similarity value. The methods are useful for classifying a breast
XX cancer patient according to prognosis. Kits and computer program products
XX are useful for data analysis using the diagnostic, prognostic and
XX statistical methods of the invention. This sequence corresponds to a
XX marker used in the method of the invention.
XX
XX Sequence 2744 BP; 786 A; 585 C; 601 G; 772 T; 0 U; 0 Other;
XX
XX
XX Alignment Scores:
XX Pred. No.: 3 139-129 Length: 2744
XX Score: 1359.50 Matches: 301
XX Percent Similarity: 40.32% Conservative: 1
XX Best Local Similarity: 40.19% Mismatches: 0
XX Query Match: 85.02% Indels: 447
XX DB: 13 Gaps: 1
XX
XX US-10-659-004-104 (1-302) X ADR25539 (1-2744)
QY 1 MetAlaArgLYsLeuSerValIleLeuIleLeuThrPheAlaLeuSerValThrAsnPro 20
Db 138 ATGCGAGAGAAATTATGTTATCTTGATCTGACCTTGGCCCTCTGTGCACAAATCCC 197
QY 21 LeuHisGluLeuLYsAlaAlaAlaPheProGlnThrThrGluLYsIleSerProAsnTrp 40
Db 198 CTTCATGAACTTAAAGAGAGCTGCTTCCACAGCCACTGAGAAATTAATGTCGAATTGG 257

```

QY 41 GluSerGlyIleAsnValAspLeuAlaIleSerThrArgIleTyrHisLeuGlnIleu 60  
 |||||  
 Db 258 GAATCTGGCATTAATTTGACTTGCACTTTCACACCGCAATATCACTCAACAGCTT 317  
 QY 61 PheTyrArgIlyGlyAsnAsnSerLeuSerValGluGlyPheArgIleLeuGln 80  
 |||||  
 Db 318 TTCTACCGCTATGGAGAAATATACTTGTGACGTGAAGGTTCAAAAATTACTTCAA 377  
 QY 81 AsnIleGlyIleAspIleValIleLeuArgIleHisGlyHisIleAspHisIleSer 100  
 |||||  
 Db 378 AATATAGCATATGATTAAGATTAAAGAAATCCATATACCATGACACACGACCATCTCA 437  
 QY 101 AspHisGluHisIleSerAspHisGluArgHisIleSerAspHisIleSerAsp 119  
 |||||  
 Db 438 GACCAAGCATCTCAAGACCATGAGCGTCACTCAGACATGACATCACTCAGACAC 497  
 QY 119 ----- 119  
 Db 498 GAGCATCACTGACCATATCATGCTGCTTGTGTAAATAAGCGAAAGCTTTTGC 557  
 QY 119 ----- 119  
 Db 558 CCAAGCATGATCAGATAGTTCAGTAAAGATCCTAGAAACAGCGGAGAAAGAGCT 617  
 QY 119 ----- 119  
 Db 618 CACCGACCAAGACATGCCAGTGTGAAAGAAATGCAAGACAGTGTAGTCTAGTGA 677  
 QY 119 ----- 119  
 Db 678 GTGACCTCACTGCTGTACAACAGCTGTCTGAGAGAACTACTTTCTAGAGAAATAG 737  
 QY 119 ----- 119  
 Db 738 ACTCAAGACCTGGAAAACTCTCCCAAGATGTAAAGACCTCACTCCACAGTGTCTC 797  
 QY 119 ----- 119  
 Db 798 ACATCAAGAGCCGGGTGAGCCGGCTGGCTGTAGGAAAACAAAATGAATCTGTGAGT 857  
 QY 119 ----- 119  
 Db 858 CCGGAAAAAGGCTTATGTATTCAGAAACAAATGAATAATCTCAGAGATGTTTCAT 917  
 QY 119 ----- 119  
 Db 918 GCATCAAGCTATGATCATCTCATGGCATGGGATCCAGTTCGGTGAATGCAACAG 977  
 QY 119 ----- 119  
 Db 978 TTCAACTATCTCTGTCCAGCATCATCAACCAAAATGATGCTAGATCTTGTGATTCAT 1037  
 QY 119 ----- 119  
 Db 1038 ACAAGTGAAGAGGCTGAATCCCTCCAAAGACTATTCATTACAAATAGCCTGGTT 1097  
 QY 119 ----- 119  
 Db 1098 GGTGTTTATAGCATTTCCATCATCAGTTTCTGCTGTCTGCGGGTATCTTAGTG 1157  
 QY 119 ----- 119  
 Db 1158 CCTCTCATGAATCGGGTGTCTTTCAAATTTCTCTGAGTTTCTTGTGGCACTGGCGGTT 1217  
 QY 119 ----- 119  
 Db 1218 GGGACTTGAAGTGATGCTTTTTCACCTTCTTCACATTCATGCAAGTCAACAC 1277  
 QY 119 ----- 119  
 Db 1278 CATAGTCATAGCATGAGAAACAGCAATGGAATGAAGAGACCACTTTTCAGTCAT 1337  
 QY 119 ----- 119

Db 1338 CTGTCTTCTCAAAACATAGAAAGAAAGTCCATTTTGAATTCACAGTGGAAAGGTCTAACA 1397  
 QY 119 ----- 119  
 Db 1398 GCTCTAGAGAGCCTGTATTTTCATGTTTCTTGTGAAACATGCTCATGATGATCAACAA 1457  
 QY 119 ----- 119  
 Db 1458 TTTAAGATAGAAAGAAAGAAATCAGAGAAACCTGAAATGATGATGTGGAGATT 1517  
 QY 119 ----- 119  
 Db 1518 AAGAAGACTGTTCAGATGATGAATCTCACTTTCAACAAATGAGAGAAAGATGATACA 1577  
 QY 119 ----- 119  
 Db 1578 GATGATGAACTGAAGGCTATTTAGAGACAGACTCAAGAGCCCTGCCACTTGATTCT 1637  
 QY 119 ----- 119  
 Db 1638 CAGAGCCTGCAGTCTTGAAAGAAAGATCATGATAGCTCATGCTCATCAAGAA 1697  
 QY 119 ----- 119  
 Db 1698 GTCTACATGAATGTATGCCAGAGGTGCAAGATTAATGCCATTCAATTTCCAGAT 1757  
 QY 119 ----- 119  
 Db 1758 ACACTCGGCGAGTACAGATCTCATTCACCAACCATCATGACTACATATTTCTCCAT 1817  
 QY 120 ----- 1817  
 Db 1818 CATCACCAACCAAAACCAACATCTCACAGTACAGCCAGCCTACTCTCGGAGAGAG 1877  
 QY 134 LeuValAspAlaGlyValAlaThrLeuAlaTrpMetValIleMetGlyAspGlyLeuHis 1937  
 |||||  
 Db 1878 CTGAAGATGCCGGCTGCGCCACTTGTGCTGATGTGATTAATGAGTATGAGCTGAC 1937  
 QY 154 AsnHisSerAspGlyLeuAlaIleGlyAlaAlaThrArgIleLeuSerSerGlyLeu 173  
 |||||  
 Db 1938 AATTTCAGCGATGCTTACGATTTGATGCTGATTTTCTGAAAGCTTATCAAGTGTTA 1997  
 QY 174 SerHisSerValAlaValPheCysHisGlyLeuProHisGlyLeuGlyAspPheAlaVal 193  
 |||||  
 Db 1998 AGTACTTCTGTTGCTGTCTGTCTGTATGATGAGTGTGCTCATGATTAAGTGTGACTTGTCTGTT 2057  
 QY 194 LeuLeuValAlaGlyMetThrValIleGlyAlaValLeuTyrAsnAlaLeuSerAlaMet 213  
 |||||  
 Db 2058 CTACTAAAGGCTGGCATGACCGTTAAGCAGGCTGTCTTATTAATGACATTTGTCAAGCATG 2117  
 QY 214 LeuAlaTyrLeuGlyMetAlaThrGlyIlePheIleGlyHisTyrAlaGluAsnValSer 233  
 |||||  
 Db 2118 CTGGCGTATCTTGGAAATGGCAACAGAAATTTTCAATTTGCTAATTTGCGAAAAATTTCT 2177  
 QY 234 MetTrpIlePheAlaLeuThrAlaGlyLeuPheMetHisValAlaLeuValAspMetVal 253  
 |||||  
 Db 2178 ATGTGATATTTGCACTTACGCTGCTGTATTCATGTATGTGCTGTGATATAGTGA 2237  
 QY 254 ProGluMetLeuHisAsnAspAlaSerAspHisGlyCysSerArgTrpGlyTyrPhePhe 273  
 |||||  
 Db 2238 CTTGAATGTGCAATAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2297  
 QY 274 LeuGlnAsnAlaGlyMetLeuLeuGlyPheGlyIleMetLeuLeuIleSerIlePheGly 293  
 |||||  
 Db 2298 TTACAGAAATGCTGGAGTCTTTTGGTGTGGAATTAATGTTACTTATTTTCATATTTGAA 2357  
 QY 294 HisIleValIleValPheArgIleAsnPhe 302  
 |||||  
 Db 2358 CATAAATCGTGTTCGTATTAATTTTC 2384  
 RESULT 8  
 ADR87269

ID ADR87269 standard; CDNA; 2744 BP.  
XX ADR87269;  
AC  
XX 04-NOV-2004 (first entry)  
DT  
XX CDNA encoding LIV-1cancer antigen.  
DE  
XX LIV-1; cancer; gene therapy; LIT; LIV-1-ZIP; Zinc transporter;  
KW cell growth; angiogenesis; cell remodeling;  
KW zinc-influx transporter protein; breast cancer; prostate cancer;  
KW colorectal carcinoma; cytotoxic; papillary carcinoma; antigen; ss.  
OS Unidentified.  
FH Key Location/Qualifiers  
FT CDS 138..2387  
FT /tag=a  
FT /product="LIV-1 Protein"  
FT  
FN WO2004067564-A2.  
PD 12-AUG-2004.  
XX  
PD 29-JAN-2004; 2004MO-US003194.  
XX  
PR 29-JAN-2003; 2003US-0443712P.  
XX  
PA (PROT-) PROTEIN DESIGN LABS INC.  
XX  
PI Law D, Gish KC, Murray R, Culp P;  
DR WPI; 2004-580975/56.  
XX P-PSDB; ADR87270.  
DR  
PT New anti-LIV-1 antibodies that specifically bind to LIV-1 proteins,  
PT useful for diagnosing, prognosing or treating cancer, particularly  
PT prostate or breast cancer.  
PT  
XX Claim 1; SEQ ID NO 1; 66pp; English.  
PS  
XX  
XX The patent discloses antibodies that specifically bind to LIV-1 protein  
XX which can be used for diagnosis, prognosis, and therapy of cancer. LIV-1  
XX is a member of the LIT (LIV-1-ZIP Zinc transporters subfamily of zinc  
XX transporter proteins. Zinc plays an essential role in cell growth and is  
XX a cofactor of over 300 enzymes, including enzymes important in  
XX angiogenesis and cell remodeling. The structure of LIV-1 implicates a  
XX role for the protein as a zinc-influx transporter protein. LIV-1  
XX distribution studies indicate primary expression in breast, prostate,  
XX pituitary gland and brain tissue. This protein has also been implicated  
XX in certain cancerous conditions like breast cancer, prostate cancer and  
XX colorectal carcinoma. The antibody used in the invention is conjugated to  
XX an effector component selected from a fluorescent label, a radioisotope  
XX or a cytotoxic chemical. The cytotoxic chemical is auristatin-E. Treating  
XX an individual with prostate or breast cancer comprises administering the  
XX above antibody or double-stranded ribonucleic acid. The antibody is a  
XX monoclonal antibody, an antibody fragment, or a humanized antibody. The  
XX invention also provides double-stranded ribonucleic acids that bind to  
XX mRNA encoded by the LIV-1 nucleic acid. These mRNA molecules can be used  
XX as a cytotoxic agent for LIV-1 expressing cells. LIV-1 has been  
XX implicated in various types of cancers including, papillary carcinoma,  
XX and squamous cell carcinoma in the bladder, ductal and lobular carcinoma  
XX in the breast, epithelial neoplasia, small cell tumours, and germ cell  
XX tumours in prostate. The sequence presented here is the CDNA encoding LIV  
XX -1 cancer antigen.  
XX  
SQ Sequence 2744 BP; 786 A; 585 C; 601 G; 772 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 3,19e-129 Length: 2744  
Score: 1359.50 Matches: 301  
Percent Similarity: 40.32% Conservative: 1  
Best Local Similarity: 40.19% Mismatches: 0

Query Match: 85.02% Indels: 447  
DB: 13 Gaps: 1  
US-10-659-004-104 (1-302) x ADR87269 (1-2744)  
QY 1 MetAlaArgLysLeuSerValIleLeuIleLeuThrPheAlaLeuSerValThrAspPro 20  
DB 138 ATGGCGAGAGATTATCTGTAATCTTATCTGACCTTTGCCCTCTGTCAAAATCCC 197  
QY 21 LeuHisGluLeuYsaIaIaIaPheProGlnThrThrgluYsIleSerProAsnTrp 40  
DB 198 CTTCACTAACTAAAGCAGCTTCCCCAGACCACTGAGAAATTAATGTCGAATGG 257  
QY 41 GluSerGlyIleAsnValaPleuAlaIleSerThrArgGlnTrHsIleGlnGluLeu 60  
DB 258 GAATCTGGCATTAATGTTGACTTGGCAATTTCCACGGCAATATCAATCAACAGCTT 317  
QY PheTyArgTrgIyGluYsaIaIaIaPheProGlnThrThrgluYsIleSerProAsnTrp 80  
DB 318 TTCTACCGCTRTGAGAAATAATTTCTTGTCACTTGAAGGTTCAAGAAATTAATCTTCA 377  
QY 81 AsnIleGlyIleAspLysIleYsaIaIaIaIleHisIleHisIleAspHisSer 100  
DB 378 AATATAGGCATAGATTAAGATTAAGAAATCCATATACCACTGACCAACATCACTCA 437  
QY 101 AspHisGluHisSerAspHisGluArgHisSerAspHisGluHisSerAsp 119  
DB 438 GACCACGACATCACTGACACCATGACGCTCACACCATGACATCACTCAACAC 497  
QY 119 ----- 119  
DB 498 GAGCATCACTGTGACCATATCATGCTGTCTGTGTAATAATAGCAAAAGCTTTTGC 557  
QY 119 ----- 119  
DB 558 CCAGACCATGACTGATAGTATTGAGTAAAGTCTTAAGAAACGACGAGGAAGAGCT 617  
QY 119 ----- 119  
DB 618 CACCGACCAAGACATGCCAGTGGTAGAAGAAATGTCAGAGAGTGAAGTGA 677  
QY 119 ----- 119  
DB 678 GTGACCTCACTGTGTAACAACACTGTCTGTAAGAAATCACTTTAGAGACAATAGAG 737  
QY 119 ----- 119  
DB 738 ACTCAAGACCTGGAAAACTTTCCTCCAAAGATGTAAAGACTCCACTCCACCAAGTGC 797  
QY 119 ----- 119  
DB 798 ACATCAAAAGACCCGGGTGAGCCGGCTGGCTGTGAAACAATGAATCTGTGAGTAGAG 857  
QY 119 ----- 119  
DB 858 CCCGAAAGAGCTTTATGTATTCAGAAACACAATGAAATCTCCAGAGTGTTCAT 917  
QY 119 ----- 119  
DB 918 GCATCAAAAGCTACAGACATCATGATGGCATCGAGTTCCGCTGAATGCAACAGAG 977  
QY 119 ----- 119  
DB 978 TTCAACTATCTCTGTCCAGCATCATCAACCAATGATGTAGTATCTTGTGATTCAT 1037  
QY 119 ----- 119  
DB 1038 ACAAGTAAAGAGAGCTGAATCCTCCAAAGACTATTCAATACAAATAGCTGGGTT 1097  
QY 119 ----- 119  
DB 1098 GGTGTTTAAATAGCATTTTCATCATCATGAGTTTCTGTCTGTGAGGGTTATCTTAAGTG 1157

```

Oy 119 ----- 119
Db 1158 CCTCTGATGAATGGGGTGTTCCTTCAAAATTTCTCTGAGTTCTTGTGGCACTGGCCGTT 1217
Oy 119 ----- 119
Db 1218 GGGACTTGAAGTGTGATGCTTTTTCACCTTTCCACATTTCTCATGCAAGTCAACAC 1277
Oy 119 ----- 119
Db 1278 CATAGTCATAGCCATGAAGAACGACCAATGAAATGAAAGAGACCACTTTTCAGTCAT 1337
Oy 119 ----- 119
Db 1338 CTGCTCTTCAAAACATAGAGAAAGGCTTATTTGATTTCCAGCTGAGGGCTTAACA 1397
Oy 119 ----- 119
Db 1398 GCTGTAGAGAGGCTGTATTTCATGTTTCTGTGTAACATGTCCTCATATTGATCAAGAA 1457
Oy 119 ----- 119
Db 1458 TTTAAAGATTAAGAGAAAAAGATGACAGAAACCTGAAAAATGATGATGATGAGAGATT 1517
Oy 119 ----- 119
Db 1518 AAGAAAGAGTTGTCAGATGATGATCTCAACTTCAACAAATGAGAGAAAGTAGATACA 1577
Oy 119 ----- 119
Db 1578 GATGATGGAAGGCTATTTAGAGAGAGCTCACAAGAGCCCTCCACCTTTGATTCT 1637
Oy 119 ----- 119
Db 1638 CAGCAGCTGCGATCTTGAGAGAGAGAGTATGATGATGATGATGATGATGATGATGATG 1697
Oy 119 ----- 119
Db 1698 GTCTACAAATGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1757
Oy 119 ----- 119
Db 1758 ACACCTGGCCAGTCAAGAGATCTCATGACCACTCATGATCATATTTCTCAT 1817
Oy 120 ----- 119
Db 1818 CATCACCAACCAACCAACCACTCTCATGATGATGATGATGATGATGATGATGATGATGATG 1877
Oy 134 LeuysaAaPAlaGlyValAlaThrLeuAlaTrpMetValIleMetGlyAaPAlaGlyLeuHis 153
Db 1878 CTGAAAGATGCCGGCGTGGCCACTTTGACCTGATGATGATGATGATGATGATGATGATGATG 1937
Oy 154 AenPheSerAaPAlaGlyValAlaThrLeuAlaTrpMetValIleMetGlyAaPAlaGlyLeu 173
Db 1938 AATTTCACGATGCGCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1997
Oy 174 SerThrSerValAlaValAlaPheCysHisGlyLeuPheProHisGlyLeuGlyAaPAlaVal 193
Db 1998 AGTACTCTGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2057
Oy 194 LeuLeuValAaGlyMetThrValIleGlyAlaValIleLeuTrpAlaIleLeuSerAlaMet 213
Db 2058 CTACTAAAGGCTGGCATACCGTTTAAAGAGCGCTGCTTTTAAAGCAATGTCACCCATG 2117
Oy 214 LeuAlaTrpLeuGlyMetAlaThrGlyIlePheIleGlyHisGlyTrpAlaGlyAaPAlaSer 233
Db 2118 CTGGGATCTTTGGAAATGGCAACAGCAATTTTCATGCTCATGCTGCAAAATTTTCT 2177
Oy 234 MetTrpIlePheAlaLeuThrAlaGlyLeuPheMetHisValAlaLeuValAaPAlaVal 253
Db 2178 ATGTGGAAATTTTGGCACTTACTGCTGCTTATTCATGATGATGATGATGATGATGATGATG 2237
Oy 254 ProGluMetLeuHisAaPAlaSerAaPAlaSerAaPAlaSerAaPAlaSerAaPAlaSerAa 273

```

```

Db 2238 CCTGAAATGCTGCACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2297
Oy 274 LeuGlnAaAaAlaGlyMetLeuGlyPheGlyIleMetLeuLeuIleSerIlePheGly 293
Db 2298 TTACGAAATGCTGGGATGCTTTTGGATTTTGGAAATGATGATGATGATGATGATGATGATG 2357
Oy 294 HisValIleValAlaPheAlaGlyLeuHis 302
Db 2358 CATAAATGCTGTTTCTGATTAATTTTC 2384

RESULT 9
ADP56127
ID ADP56127 standard; cDNA, 2744 BP.
AC ADP56127;
DT 18-NOV-2004 (first entry)
DE Human PRO cDNA sequence SEQ ID NO:2103.
KW human; PRO; immune related disease; inflammatory immune response;
KW immune response stimulation; antiallergic; antianemic; antiarthritic;
KW antirheumatic; antidiabetic; antiinflammatory; antipneumatic;
KW antirheumatic; antihypoid; CNS; dermatological; gastrointestinal;
KW haemostatic; hepatocytic; immunostimulant; immunosuppressive; muscular;
KW nephrotropic; neuroprotective; osteopathic; respiratory; vasodilatoric;
KW virucide; gene therapy; gene; ss.
OS Homo sapiens.
PN WO2004039956-A2.
PD 13-MAY-2004.
XX 28-OCT-2003; 2003WO-US034381.
XX 29-OCT-2002; 2002US-0422472P.
XX (GETH ) GENENTECH INC.
XX Aggarwal S, Clark H, Gurney AL, Schoenfeld J, Williams PM;
XX Wood WJ, Wu TD;
XX WPI; 2004-376182/35.
XX P-PSDB; ADP56128.
XX

PT New PRO polynucleotides and polypeptides, useful in useful in diagnosing
PT and treating an immune related disease, e.g. systemic lupus
PT erythematosus, rheumatoid arthritis, diabetes mellitus or asthma and in
PT stimulating an immune response.
XX
XX
XX Claim 2; SEQ ID NO 2103; 3009pp; English.
XX
XX The present invention describes an isolated PRO nucleic acid (1). Also
XX described: (1) a vector comprising (1); (2) a host cell comprising the
XX vector of (1); (3) a process for producing a PRO polypeptides; (4) an
XX isolated PRO polypeptide; (5) a chimeric molecule comprising the
XX polypeptide of (4) fused to a heterologous amino acid sequence; (6) an
XX antibody which specifically binds to a polypeptide of (4); (7) a
XX composition of matter comprising a polypeptide of (4), an agonist or
XX antagonist of the polypeptide or an antibody that binds to the
XX polypeptide in combination with a carrier; (8) an article of manufacture
XX comprising a container, a label on the container and a composition of
XX matter of (7); (9) a method of treating an immune related disease in a
XX mammal; (10) a method for determining the presence of a PRO polypeptide
XX in a sample suspected of having the polypeptide; (11) a method of
XX diagnosing an immune related disease or an inflammatory immune response
XX in mammal; (12) a method of identifying a compound that inhibits or
XX mimics the activity of or expression of a gene encoding a PRO polypeptide
XX; and (13) a method of stimulating the immune response in a mammal. The
XX PRO sequences have antiallergic, antianemic, antiarthritic,
XX antirheumatic, antidiabetic, antiinflammatory, antipneumatic,

```

CC antineuritic, antithyroid, CNS, dermatological, gastrointestinal,  
CC haemostatic, hepatotropic, immunostimulant, immunosuppressive, muscular  
CC neurotropic, neuroprotective, osteopathic, respiratory, vasotropic and  
CC vinculin activities, and can be used in gene therapy. The nucleic acid  
CC (I) and the encoded polypeptides, compositions, kits and methods are  
CC useful in diagnosing and treating an immune related disease and in  
CC stimulating an immune response. The present sequence represents a human  
CC PRO nucleotide sequence from the present invention.

Sequence 2744 BP; 786 A; 585 C; 601 G; 772 T; 0 U; 0 Other;

### Alignment Scores:

Pred. No.:	3,198-129	Length:	2744
Score:	1359.50	Matches:	301
Percent Similarity:	40.328	Conservative:	1
Best Local Similarity:	40.194	Mismatches:	0
Query Match:	85.028	Indels:	447
DB:	13	Gaps:	1

US-10-659-004-104 (1-302) X ADP56127 (1-2744)

Qy	1	MetAlaArgyLeuSerValIleLeuIleLeuThrPheAlaLeuSerValThrAspPro	20
Db	138	ATGGCGAGAGATTATCTGTATCTTGATCTGCACCTTGGCCCTCTGTGCACAAATCC	197
Qy	21	LeuHi6GIuLeuLySaIaaIaaIaPheProGInThrTrGIuLySiLeSerProAsnTrp	40
Db	198	CTTCATGAACTAAAGACAGCTGCTTTCCCCGACGACCATGAGAAATTTAGTCCGAATTGG	255
Qy	41	GIuSerGIyIleAsnValAspLeuAlaIleSerThrArgGInTrpHisLeuGInGluLeu	60
Db	258	GAATCTGGCAATTATATGTGTGACTTGGCAATTTCCACACGGCAATATCATCTCAACACTT	317
Qy	61	PheTyArgTrpGIyGIuGluAsnSerLeuSerValGIuGIyPheArgLyLeuLeuGIn	80
Db	318	TTCTACCGCTATGGAGAAATATTTCTTTGTGCAGTTGAAGGGTTCAGAAATTACTTCAA	377
Qy	81	AsnIleGIyIleAspLySileuYsArgIleHisIleHisAspHisAspHisHisSer	100
Db	378	AATATGGCATTAATTAAGATTAAAGAAATCCATATACACATGACACAGACCATCTCA	437
Qy	101	AspHisGIuHisHisSerAspHisGIuArgHisSerAspHisGIuHisHisSerAsp--	119
Db	438	GACCACAGCATCATCTCAAGCATGAGCTCACTCAGACATGAGCATCTCAGACACAC	497
Qy	119	-----	119
Db	498	GAGCATCATCTGACCATATATCATGCTGCTTGTGTAAAAATAGCGAAAAGCTTTTGC	557
Qy	119	-----	119
Db	558	CCAGACCATGACTCATAGTAGTTCAGTAAAGATCTCTAGAAAACGCCAGGGAAAAGAGCT	617
Qy	119	-----	119
Db	618	CACCGACCAAGACATGCGAGTGGTAGAAGAAATGTCAGAGACAGTGTAGTGTAGAA	677
Qy	119	-----	119
Db	678	GTAACCTCAACTGTGTACAACACTGTCTCTGAGAGAACTCACTTTCTAGAGACAATAGAG	737
Qy	119	-----	119
Db	738	ACTCCAAAGACTGGAAAACTTTCCCCAAAGATGTAAAGACAGCTCCACCTCCACCCAGTGC	797
Qy	119	-----	119
Db	798	ACATCAAAAGAGCCGGGTGAGCCGGCTGGCTGGTAGAAAAACAATGAATCTGTGAGTAGAG	857
Qy	119	-----	119
Db	858	CCCCGAAAAGCTTTATGTATTCAGAAAACAATAATGAATCTCTCAGAGATGTTTCAAT	917

QY	119	-----	119
Db	918	GCATCAAAAGCTACTGACATCTCATGGCATGGGCATCCAGGTTCCGCTGAATGCAACAGAG	977
QY	119	-----	119
Db	978	TTCAACTATCTGTGCCAGCATGATCAACCAATGTAGTGTAGTCTTGTCTGATTGAT	1031
QY	119	-----	119
Db	1038	ACAAGTGAAGAAAGGCTGAATCCCTCCAAAGACTATTGATTACAATAGCCTGGGTT	1097
QY	119	-----	119
Db	1098	GGTGTTTATAGCATTTCCATCATGAGTTTCCGTCTCTGCTGGGGTTATCTTAGTG	1153
QY	119	-----	119
Db	1158	CCTCTCATGAATCGGGTGTGTTTTCAAAATTTCTCCTGAGTTTCTTGTGGCACTGGCGGTT	1217
QY	119	-----	119
Db	1218	GGGACTTTGAGTGGTATGCTTTTTCACACTTCTTCCACATTTCTCATGCAAGTCACCAC	1277
QY	119	-----	119
Db	1278	CATAGTCATAGCATGAAGAACACAGCATGGAATGAAGAAAGAGCACCTTTCAGTCAT	1333
QY	119	-----	119
Db	1338	CTGTCTTCTCAAAACATGAGAAAGTGCTATTTTGATGTCACGTGAGGGTCTTACA	1397
QY	119	-----	119
Db	1398	GCTGTAGAGGCTGTATTTGATGTTGTTTCTTGTGAACATGTCCTCATGTGATCAACAA	1457
QY	119	-----	119
Db	1458	TTTAAAGATGAAGAAAAAGAAATCAGAGAAACCTGAAATGATGATGTGAGATT	1517
QY	119	-----	119
Db	1518	AAGAAAGCATGTTGCCAAGTAGATGAATCTTCACCTTTCACAAATGAGAGAAAGTAGATACA	1577
QY	119	-----	119
Db	1578	GATGATCGAAGCTAAGGCTATTTTACGACAGACTCACAAAGAGCCCTCCACTTTGATTCT	1637
QY	119	-----	119
Db	1638	CAGAGCCTCGAGTCTTGGAAGAAAGAGGTGATGATGCTCATGCTCATCCACAGGAA	1697
QY	119	-----	119
Db	1698	GTCACATGAATATGTATACCCAGAGGTGCAAGATTAATGCCATTGCATTCCACAGAT	1757
QY	119	-----	119
Db	1758	ACACTCGGCGCAGTCAGACGATCTCATCCACCACATGATGTAACAATATTCTCGAT	1817
QY	120	-----	133
Db	1818	CATCAACCAACCAAAACCAACCACTCTCACAGTACACGCCCTACTCTCGGGAAGAG	1877
QY	134	LeuIysApR]agIyVal]aIaThrLeuAlaTrpMetVal]I]ewetGIyAspGIyLeuHi	153
Db	1878	CTGAAGAATGCGCGCTGCGCACCTTTGSCCTCGATGATGAATGAAGGATGAGGCTCGAC	1937
QY	154	AsnPheSerAspGIyLeuAlaI]egIyAla]aIaPheThrGIuGIyLeuSerSerGIyLeu	173
Db	1938	AAITTCAGCGAGTGCCTTAGCATTTGGTGTGCTTTTACTGAAAGCTTATCAAGTGGTTTA	1997
QY	174	SerThrSerVala]aVal]PheCyHisGIuLeuProHisGIuLeuGIyAspPheAlaVal]	193

Db 1998 AGTACTTCTGTTGCTGTGTGTGTATGATGCTGCTCAATGATGATGATGCTGCTT 2057  
 Qy 194 LeuLeuValAlaGlyMetThrValValValValValValValValValValValVal 213  
 Db 2058 CTACTAAAGGGCGCATGACCTTAAAGCAGCGCTGCTTAAAGCATTTGACAGCCATG 2117  
 Qy 214 LeuAlaValLeuGlyMetAlaThrGlyIlePheIleGlyHisIleValAlaGlyAsnValSer 233  
 Db 2118 CTGGGCTATCTTGGAAATGGCAACAGAAATTTTCATTTGCTGATTAAGTGAATATTTCT 2177  
 Qy 234 MetTrpIlePheAlaLeuThrAlaGlyLeuPheMetHisValAlaLeuValAspMetVal 253  
 Db 2178 ATGTGATATTTGCACTTACGCTGCTTATTCATGTAATGTTGCTGCTGATATGATA 2237  
 Qy 254 ProGluMetLeuHisAsnAspAlaSerAspHisGlyCysSerArgTrpGlyTrpPhePhe 273  
 Db 2238 CCTGAAAGCTGTCACAAATGATGCTAGTACCATGATGATGATGATGATGATGATGAT 2297  
 Qy 274 LeuGlnAsnAlaGlyMetLeuLeuGlyPheGlyIleMetLeuLeuIleSerIlePheGlu 293  
 Db 2298 TTACGAATGCTGGGATGCTTTGGCTTTGGAATTAATTAATTAATTAATTAATTAAT 2357  
 Qy 294 HisValIleValPheArgIleAsnAsp 302  
 Db 2358 CATAAATCGTCTTTCGTAATTAATTTTC 2384

RESULT 10  
 ADR66497  
 ID ADR66497 standard; DNA; 2744 BP.  
 AC ADR66497;  
 XX 02-DEC-2004 (first entry)  
 DT  
 XX Human prostatic carcinoma derived DNA SEQ ID #3.  
 DE  
 XX human; cyrostatic; diagnosis; prostatic cancer;  
 KM differential expression analysis; ds.  
 XX Homo sapiens.  
 OS  
 XX MO2004076614-A2.  
 PN  
 XX 10-SEP-2004  
 PD  
 XX 22-FEB-2004; 2004MO-DE000433.  
 PF  
 XX 27-FEB-2003; 2003DE-01009985.  
 PR 14-MAY-2003; 2003DE-01022134.  
 RR  
 XX (HINZ/) HINZMANN B.  
 PA (DAHL/) DAHL E.  
 PA (ROSE/) ROSENTHAL A.  
 PA (HERM/) HERMANN K.  
 PA (PILA/) PILARSKY C.  
 XX  
 PI Hinzmann B, Dahl E, Rosenthal A, Hermann K, Pilarsky C, Specht T;  
 PI Schmitt A, Beckmann G, Bruemendorf T, Kinneemann H, Roepcke S;  
 PI Kinzhong L, Staub E,  
 XX WPI; 2004-653386/63.  
 DR  
 XX New nucleic acids, and encoded proteins, from prostatic cancer tissue,  
 PT useful for diagnosis, treatment and in screening for specific binding  
 PT agents.  
 XX  
 PS Claim 1; Page 1130; 1607pp; German.  
 XX  
 CC This invention describes novel cyrostatic polynucleotide and polypeptide  
 CC sequences which can be used in a method for diagnosing prostatic cancer  
 CC or the risk of developing prostatic cancer. Diagnosis is based on  
 CC determining over transcription or over expression of the sequences in

CC prostatic tissue. Screening for inhibitors of the sequences or detection  
 CC substances involves a binding assay, any compounds that bind are  
 CC selected, optionally after deconvolution of mixtures. Detection of a  
 CC predetermined minimum level of the reporter indicates the presence of  
 CC tumour cells. Inhibitors can be chosen from antisense oligonucleotides,  
 CC short-interfering RNA or ribozymes; an organic molecule of molecular  
 CC weight below 5000, preferably 300, that binds to the polypeptide; an  
 CC aptamer against the polypeptide; a (monoclonal) antibody (Ab) against the  
 CC polypeptide, preferably humanised or human; an anti-idiotypic, non-human  
 CC (monoclonal) antibody directed against Ab or any of the above derivatised  
 CC with a reporter group, cell toxin, immunostimulatory molecules and/or  
 CC radioisotope. The polynucleotides are identified in human prostatic  
 CC cancer by differential expression analysis, using DNA microarrays,  
 CC between normal and tumorous tissues, with (over)expression being detected  
 CC by quantitative PCR. Analysis of prostatic cancer samples showed that  
 CC CD24 was upregulated in many of them. Sections of tissue, isolated from  
 CC prostatic cancer patients, or subjects at risk, were incubated  
 CC sequentially with anti-human CD4 murine monoclonal antibodies;  
 CC biotinylated second antibody; streptavidin-conjugated horseradish  
 CC peroxidase and then diaminobenzidine as colour former (brown). The  
 CC samples were counterstained with hemalum (blue). Malignant cells stained  
 CC strongly but non-malignant cells only weakly. In 15 of 63 samples of  
 CC adenocarcinoma, membrane and cytoplasmic staining was very strong, and  
 CC lymph node metastases were also stained. ADR65805-ADR66954 represent the  
 CC polynucleotide and polypeptide sequences used in the method of the  
 CC invention.

SO Sequence 2744 BP; 786 A; 585 C; 601 G; 772 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 3,19e-129 Length: 2744  
 Score: 1359.50 Matches: 301  
 Percent Similarity: 40.32% Conservative: 1  
 Best Local Similarity: 40.19% Mismatches: 0  
 Query Match: 85.02% Indels: 447  
 DB: 13 Gaps: 1

US-10-659-004-104 (1-302) x ADR66497 (1-2744)

Qy 1 MetAlaArgLysLeuSerValIleLeuIleThrPheAlaLeuSerValThrAsnPro 20  
 Db 138 ATGGGAGGAAGTATATCTGTAATCTTGATCTTGACCTTGGCCCTCTGTGCAAAATGCC 197  
 Qy 21 LeuHisGluLeuValAlaAlaPheProGlnThrThrGluValIleSerProAsnTrp 40  
 Db 198 CTTGATGAATCAAAAGACGCTGCTTCCCGACGACCACTGAGAAATTAATGCGAATTGG 257  
 Qy 41 GluSerGlyIleAsnValAspLeuValIleSerThrArgGlnTrpHisGluGlnLeu 60  
 Db 258 GAATCTGGCAATTAATGTTGATGGCAATTTCCACAGCGCAATATCAATCAACAGCTT 317  
 Qy 61 PheTyraGlyTrpGlyLysAsnAsnSerLeuSerValGluGlyPheArgLysLeuGln 80  
 Db 318 TTCTACCGCTATGAGAAATTAATTTCTTGCACTTGAAGGTTCAAGAAATTAATTTCTCA 377  
 Qy 81 AsnIleGlyIleAspValIleLysArgIleHisIleHisAspHisAspHisSer 100  
 Db 378 AATATAGGCAATAGATTAAGATTAAGAAATCAATTAACCATGACCAACCATCACTCA 437  
 Qy 101 AspHisGluHisHisSerAspHisGluArgHisSerAspHisGluHisSerAsp 119  
 Db 438 GACCACGACATCACTGACGACATGACGTCATCAACCATGACATCACTGACACAC 497  
 Qy 119 ----- 119  
 Db 498 GAGCATCACTGACCATTAATCATGCTGCTTCTGTAAATAAGCAAAAGCTTTTGC 557  
 Qy 119 ----- 119  
 Db 558 CGAGACATGACTCAGATAGTTAGTAAAGATCTAGAAAACAGCCAGGGAAGAGACT 617  
 Qy 119 ----- 119

Db 618 CACGACCAAGACATGCCAGTGTGTAAGAAAGATGTCAGACAGTGTAGTCTAGTGA 677  
 QY 119 ----- 119  
 Db 678 GTGACCTGACGTGTACACACGTGTCTGTAGAGAACTCACTTTCTAGAGCAATAGAG 737  
 QY 119 ----- 119  
 Db 738 ACTCCAGACCTGGAATACTTTCCCAAGATGTAGACGCTCCACTCCACCCAGTGTG 797  
 QY 119 ----- 119  
 Db 798 ACATCAAGAGCCGGGTGAGCCGGCTGGCTGTAGAGAAACAAATGAATCTGTAGTGA 857  
 QY 119 ----- 119  
 Db 858 CCCCAGAAAGCTTTATGTATTCAGAAAACAATAATGAAATCTCAGAGTGTTCAT 917  
 QY 119 ----- 119  
 Db 918 GCATCAAGCTACTGACATCTCATGCGCATGGGCATCCAGTTCCGCTGAATGCAACAG 977  
 QY 119 ----- 119  
 Db 978 TTCAACTATCTGTCCAGCCCATCATCAACCAATGTAGTATCTTGTCTGATTCAT 1037  
 QY 119 ----- 119  
 Db 1038 ACAAGTGAAGAGGCTGAAATCCCTCCAAAGACTTATCATTAATAAGCTGGGT 1097  
 QY 119 ----- 119  
 Db 1098 GGTGTTTTATAGCAATTTCCATCATCACTTCTGTCTGTGGGGTATCTTAAGTG 1157  
 QY 119 ----- 119  
 Db 1158 CCTTCAGAAATCGGGTGTTCCTCAATTTCTCTGAGTTCTCTGTGCACTGGCCGT 1217  
 QY 119 ----- 119  
 Db 1218 GGGACTTTGAGTGTGATGCTTTTTCACACTTTCACATTCATGACAGTCAAC 1277  
 QY 119 ----- 119  
 Db 1278 CATAGTCATAGCCATGAGAACACAGCATGAAATGAAAAGAGACCACTTTCACTCAT 1337  
 QY 119 ----- 119  
 Db 1338 CTGTCTTCAAAACATAGAAAGATGCTATTTTGAATTCACGTCGAAAGGTCTAACA 1397  
 QY 119 ----- 119  
 Db 1398 GCTTAGAGAGCCTGATTTTCATGTTTCTGTGGAACATGCTCCATTTGATCAACAA 1457  
 QY 119 ----- 119  
 Db 1458 TTTAAAGATGAAGAAAAAAGATCAGAAACCTGAAATGATGATGTGGAGATT 1517  
 QY 119 ----- 119  
 Db 1518 AAGAGCAATGTGCAAGTATGATCTCAACTTTCACAAATGAGAGAAAAGTATGATCA 1577  
 QY 119 ----- 119  
 Db 1578 GATGATGAAGTGAAGCTATTTACGAGACACTCAAGAGCCCTCCACTTTGATTTCT 1637  
 QY 119 ----- 119  
 Db 1638 CAGCAGCTGAGTCTTGAGAAAGAGTCAAGTCACTGATCTCATCCACAGGAA 1697  
 QY 119 ----- 119  
 Db 1698 GTCTCAATGAATATGTACCAGAGGGTGCAGAAATAAATGCCATTCACTTTCCAGAT 1757

QY 119 ----- 119  
 Db 1758 AACTCGGCGAGTCAGACGATCTCATTCACACCATATGATCAATCATATTTCTCAT 1817  
 QY 120 ----- 119  
 Db 1818 CATACCAACCAACCAACCAACCAATCTTACACAGTACAGCCAGGCTCTCTCGGAGAG 1877  
 QY 134 LeuLysAspAlaGlyValAlaThrLeuAlaTrpMetValIleMetGlyAspGlyLeuHis 153  
 Db 1878 CTGAAGATGCCGCGCTGCGCACCTTGGCTGTGATGAATGAGGTGATGCGCTGCAC 1937  
 QY 154 AsnProSerAspGlyLeuAlaIleGlyValAlaIlePheThrGlyLeuSerSerGlyLeu 173  
 Db 1938 AATTCAGCGATGGCTAGCAATGTGGTGTCTTTACTGAGGCTTATCAAGTGTGTTA 1997  
 QY 174 SerThrSerValAlaValAlaPheCysHisGlyLeuProHisGlyLeuGlyAspPheAlaVal 193  
 Db 1998 AGTACTTCTGTGCTGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 2057  
 QY 194 LeuLeuLysAlaGlyMetThrValIleGlyAlaValLeuTyraAlaLeuSerAlaMet 213  
 Db 2058 CTACTAAAGCTGGCATGACCCGTTAAGCAGGCTGCTCTTATTAATGATGATTCAGCCATG 2117  
 QY 214 LeuAlaTyrlLeuGlyMetAlaThrGlyIlePheIleGlyHisTyraAlaGluAsnValSer 233  
 Db 2118 CTGGCGTATCTTGGAAATGGCAACGAAATTTCTATGCTCATTAAGCTGAAAATGTTCT 2177  
 QY 234 MetTrpIlePheAlaLeuThrAlaGlyLeuPheMetHisValAlaLeuValAspMetVal 253  
 Db 2178 ATGTGATATTTGCACTTACCTGCTGCTTATTCATGATATGCTGCTGTTGATATGTA 2237  
 QY 254 ProGluMetLeuHisAsnAspAlaSerAspHisGlyCysSerArgTrpGlyTyrPhePhe 273  
 Db 2238 CCGAAATGCTGCACAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2297  
 QY 274 LeuGlnAsnAlaGlyMetLeuGlyPheGlyIleMetLeuLeuIleSerIlePheGlu 293  
 Db 2298 TTACGAATGCTGGAGTCTTTGGGTTTGGAAATGATGATGATGATGATGATGATGAT 2357  
 QY 294 HisLysIleValAlaPheArgIleAsnPhe 302  
 Db 2358 CATAAATCGTGTTCGATTAATTTTC 2384  
 Db  
 RESULT 11  
 ADR6155  
 ID ADR6155 standard; DNA; 2744 BP.  
 XX  
 AC ADR6155;  
 XX  
 DT 02-DEC-2004 (first entry)  
 XX  
 DE Human prostatic carcinoma derived DNA SEQ ID 9 #2.  
 XX  
 KW human, cytosolic; diagnosis; prostatic cancer;  
 KW differential expression analysis; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W02004076614-A2.  
 XX  
 PD 10-SEP-2004.  
 XX  
 PF 22-FEB-2004; 2004WO-DE000433.  
 XX  
 PR 27-FEB-2003; 2003DE-01009985.  
 PR 14-MAY-2003; 2003DE-01022134.  
 XX  
 PA (HINZ/) HINZMANN B.  
 PA (DAHL/) DAHL E.  
 PA (ROSE/) ROSENTHAL A.  
 PA (HERM/) HERMANN K.





[illegible]

QY	119	-----	119
Db	1518	AAGAAAGAGTTGTCAGAGTAGAATCTCAATTTCAACAATAGAGAGAAAGTAGATACA	157
QY	119	-----	119
Db	1578	GATGATCGAATCGAAAGGCTATTTCAGAGCAGACTCAAGAAGCCCTCCACTTTGATTCT	163
QY	119	-----	119
Db	1638	CAGCAGCCCTGACAGTCTTGAGAGAGAGAGATGATAGTCTCATGCTCATCCACAGAA	169
QY	119	-----	119
Db	1698	GCTTACAAATGATATGTACCCAGAGGGTCAGAAATTAATGCCATTTCAGATTCCAGAT	175
QY	119	-----	119
Db	1758	ACACTGGCCAGTCAGAGATCTCATCCACCATCATGATTCATCATATATTCTCCAT	181
QY	120	-----HisHisProHisSerHisSerGlnArgTyrSerArgGlnGlu	133
Db	1818	CATCACACACACCAAAACACCATCTCTCAGCTCAGCCAGCCAGCCGCTACTCTGGGAGGAG	187
QY	134	LeuLysAspAlaGlyValAlaThrLeuAlaTrpMetValIleMetGlyAspArgLysLeuHis	153
Db	1878	CTGAAGAAGATGCCGGGCGTCGCCACTCTGGCTGGATGGTATATAGGTATGATGCCCTGCAC	193
QY	154	AsnPheSerAspGlyLeuAlaIleGlyAlaAlaPheThrGlnGlyLeuSerSerGlyLeu	173
Db	1938	AATTTCAGCGATGCCCTAGCAATTGGTGTCTTTTACTGAAGGCTTATCAAGTGTTHA	199
QY	174	SerThrSerValAlaValAlaPheCysHisGluLeuProHisGluLeuGlyAspPheAlaVal	193
Db	1998	AGTACTTCTGTGCTGTCTGTCTGTCTGTCATGATGCTCTCATGAATTAAGTGACTTTGCTGTT	205
QY	194	LeuLeuValAlaGlyMetThrValLysGlnAlaValLeuTyrAspAlaLeuSerAlaMet	213
Db	2058	CTACTAAAGGCTGGGATGACCGTTAAGCAGGCTGTCTTTATTAATGACATTGTCCAGCCATG	211
QY	214	LeuAlaTyrLeuGlnMetAlaThrGlyIlePheIleGlyHisTyrAlaGluAsnValSer	213
Db	2118	CTGGGCTATCTTGGAAATGGCAACAGAAATTTTCATCTGTCATTAAGCTGAAATGTTTCT	217
QY	234	MetTrpIlePheAlaLeuThrAlaGlyLeuPheMetHisValAlaLeuValAspMetVal	253
Db	2178	ATGTGATATATTGCACTACTCTGCTGCTTATTCATGATATGTTGCTCTGGTATATAGTGA	223
QY	254	ProGluMetLeuHisAsnAspAlaSerAspHisGlyCysSerArgTrpGlyTyrPhePhe	273
Db	2238	CCTGAATAGCTGCACAAATGATGCTAGTACCACTGATGTAAGCCGCTGGGGGATTTTCTTT	229
QY	274	LeuGlnAsnAlaGlyMetLeuLeuGlyPheGlyIleMetLeuLeuIleSerIlePheGlu	293
Db	2238	TTACGAATGCTGGGATGCTTTTGGGCTTTTGGAAATTAATTAATTAATTTCCATATTGAA	235
QY	294	HisLysIleValPheArgIleAsnDhe	302
Db	2358	CATAAATCGTGTTCTGATATAAATTTTC	2384
RESULT 13			
ACN88857			
ID ACN88857 standard; DNA, 3537 BP.			
XX ACN88857;			
XX AC			
XX DT 02-DEC-2004 (first entry)			
XX DE Breast cancer related marker, seq id 10007.			
XX XX Cancer; breast; tumour; cytostatic; marker; detection; therapy; ds.			
XX			

Breast cancer related marker, seq id 10007.  
Cancer; breast; tumour; cytostatic; marker; detection; therapy; ds

OS Homo sapiens.  
XX US200309974-A1.  
XX 29-MAY-2003.  
XX 18-JUL-2002; 2002US-00198846.  
XX 18-JUL-2001; 2001US-0306220P.  
XX (MILL-) MILLENNIUM PHARM INC.  
XX Lillie J, Xu Y, Wang Y, Steinmann K;  
XX WPI; 2003-787014/74.  
XX  
XX Novel isolated polypeptide associated with breast cancer, useful for  
XX detecting presence of polypeptide in sample, as a marker for breast  
XX cancer.  
XX  
XX Disclosure; SEQ ID NO 10007; 36pp; English.  
XX  
XX The invention relates to an isolated polypeptide (I) associated with  
XX breast cancer which is encoded by a nucleic acid molecule comprising a  
XX nucleotide sequence (SI). Further disclosed is an antibody that binds to  
XX the polypeptide of the invention. The activity of the polypeptide of the  
XX invention may be described as cytostatic. The antibody is useful for  
XX detecting the presence of (I) in a sample. Nucleic acid molecules of the  
XX invention are useful in the detection of breast tumors. (I) is useful as  
XX a marker for breast cancer and in breast cancer therapy. Sequences given  
XX in records ACN78851-ACN92934 represent nucleic acid markers associated  
XX with breast cancer. Note: The sequence listing does not form part of the  
XX specification but may be obtained in electronic format from the USPTO web  
XX site at [seqdata.uspto.gov/sequence.html?docid=2003099974](http://seqdata.uspto.gov/sequence.html?docid=2003099974)  
XX  
XX  
XX Sequence 3537 BP; 1060 A; 707 C; 747 G; 1016 T; 0 U; 7 Other;  
XX  
XX  
XX Alignment Scores:  
XX Pred. No.: 4,516-129 Length: 3537  
XX Score: 1359.50 Matches: 301  
XX Percent Similarity: 40.32% Conservative: 1  
XX Best Local Similarity: 40.19% Mismatches: 0  
XX Query Match: 85.02% Indels: 447  
XX DB: 11 Gaps: 1  
XX  
XX US-10-659-004-104 (1-302) x ACN88857 (1-3537)  
XX  
XX QY 1 MetAlaArgLysLeuSerValIleLeuIleLeuThrPheAlaLeuSerValThrAsnPro 20  
XX DB 138 ATGGCGAGGAGTATCTGTATCTTGATCTGACCTTGGCCCTCTCTGTCAAAATCC 197  
XX QY 21 LeuHISGLuLeuAlaAlaAlaPheProGlnThrGluLysIleSerProAsnTrp 40  
XX DB 198 CTTCATGAACTAAAGCAGCTCTTCCCGCAGACCTGAGAAATTAATGTCGATGG 257  
XX QY 41 GluSerGlyIleAsnValAspLeuAlaIleSerThrArgLysIleGlnGlnLeu 60  
XX DB 258 GAATCTGGCATTAATCTTGCATTGGCAATTTCCACCGGCATTAATCACTACAAAGCTT 317  
XX QY 61 PheTyrArgTyrGlyGluAsnAsnSerLeuSerValGluGlyPheArgLysLeuGln 80  
XX DB 318 TTCCTACCGCTAGAGAAAATAATCTTTGTCAATTGAAGGGTTCAGAAAATTACTTCA 377  
XX QY 81 AsnIleGlyIleAspLysIleLeuArgGlnIleHisIleHisAspHisAspHisSer 100  
XX DB 378 AATAATAGGCGATGATTAAGATTAAGAAATCCATATACACCATGACACGACCATCATCA 437  
XX QY 101 AspHisGluHisHisSerAspHisGluArgHisSerAspHisGluHisHisSerAsp 119  
XX DB 438 GACCGAGGAGCATCTCAGACCATGAGCTCCTCAGACCATGAGCATCATCTCAGACCA 497  
XX QY 119 ----- 119

DB 498 GAGCATCACTCTGACCAATATCATATGCTCTTCTGTAAATAAGCAAAAGCTTTTGC 557  
QY 119 ----- 119  
DB 558 CCAGACCATGACTGATAGATTGAGTAAGATCTTAGAAAACAGCAGGGAGAGAGCT 617  
QY 119 ----- 119  
DB 618 CACCGACCAAGACATGCGAGTGTAGAAAGAAATGTCAAGACAGTGTAGTGTAGAA 677  
QY 119 ----- 119  
DB 678 GTGACTCACTGTGTACAACACTGTCTGTGAAGAACTCACTTTCTAGAGACAATAGAG 737  
QY 119 ----- 119  
DB 738 ACTCCAGACCTGGAAAACTTTCCCAAGATGTAGACGCTCAGTCCAGCAGTGTG 797  
QY 119 ----- 119  
DB 798 ACATCAAGACCGGGGTGAGCCGGCTGGTGTAGAAAACAATGAATCTGTGATGAG 857  
QY 119 ----- 119  
DB 858 CCGGAAAAGCTTTATGTAATTCAGAAACAAATGAAATCTCAGAGTGTTCAT 917  
QY 119 ----- 119  
DB 918 GCATCAAGACTACTGACATCTCATGAGATGGCATCGAGTTCCGCTGAATGACAAGAG 977  
QY 119 ----- 119  
DB 978 TTCAACTATCTCTGTCCAGCATCATCAACCAATTGATGTAGTATCTGTCTATTCAT 1037  
QY 119 ----- 119  
DB 1038 ACAAGTAAAGAGGCTGAATCTCTCAAGACTATTCATTAACAATAGCCTGGGT 1097  
QY 119 ----- 119  
DB 1098 GGTGTTTATAGCATTTCCATCATCATGATTTCTGTCTGTGCGGGTATCTTAGTG 1157  
QY 119 ----- 119  
DB 1158 CCTCTCATGAATCGGGTGTTCATTAATTTCTCTGAGTTCTTGGGACTGGCGGT 1217  
QY 119 ----- 119  
DB 1218 GGACATTTGAGTGTATGCTTTTACACCTTCTCCACATTTCTCATGACAGTCAAC 1277  
QY 119 ----- 119  
DB 1278 CATAGCATAGCCATGAGAAACAGCAATGAAATGAAGAGAGCACTTTTCAATCAT 1337  
QY 119 ----- 119  
DB 1338 CTGTCTTCTCAAAATAGAAAGAAAGTCCATTTTGTGATTTCCAGTGAAGGCTTAACA 1397  
QY 119 ----- 119  
DB 1398 GCTCTAGAGGCTGTATTTTCATGTTCTTTGTGACATGCTCATATGATGATCAACAA 1457  
QY 119 ----- 119  
DB 1458 TTTAAGATTAAGAAAGAAAGAAATCAGAGAAACCTGAAAAATGATGATGTGAGATT 1517  
QY 119 ----- 119  
DB 1518 AAGAGAGAGTTGTCCAGTATGATCTCACTTTCAACAATAGAGAAAGTAGATACA 1577  
QY 119 ----- 119  
DB 1578 GATGATGAAGTGAAGCTATTATAGAGAGAGATCACAAGAGCCCTCCACTTTGATTTCT 1637



Db 361 GAGCATCACTGACCATGATCATCATCCACATATCATGCTTCTGTGTAATAAT 420  
 QY 119 ----- 119  
 Db 421 AAGGMAAAGCTCTTGGCCAGACCATGATCAGATAGTGAAGAAATCTAGAAAC 480  
 QY 119 ----- 119  
 Db 481 AGCCAGGGGAAAGAGCTCACCGACCAAGAAATGCCAGTGTAGAAAGATGTCAAGAC 540  
 QY 119 ----- 119  
 Db 541 AGTGTAGTGTAGGAAGTGAACCTCACTGTGTACAACGTCTCTGAAGAACTGAC 600  
 QY 119 ----- 119  
 Db 601 TTTCTAGAGACAAATGAGACTCCAGACCTGGAAAACTTTCCCAAGATGTAAGCAGC 660  
 QY 119 ----- 119  
 Db 661 TCCACTCCACCCAGTGTCACTCAAAAGCCGGGTGAGCCGGCTGGTGTAGGAAAAACA 720  
 QY 119 ----- 119  
 Db 721 AATGAATCTGTAGTAGAGCCCGAAAAAGGCTTATGTATTCAGAAACAAATGAAAT 780  
 QY 119 ----- 119  
 Db 781 CCTCAGAGTGTTCATGATCAAAAGCTACTGATCTCATGCGATGGGATCCAGGTT 840  
 QY 119 ----- 119  
 Db 841 CCGCTGAATGCAACAGATTCAACTATCTGTCTCAGCATCATCAACCAATTGATGCT 900  
 QY 119 ----- 119  
 Db 901 AGATCTGTCTGATTCATACAGTAGAAGAAAGGCTGAATCCCTCAAGACCTATTCA 960  
 QY 119 ----- 119  
 Db 961 TTCAAAATAGCCTGGGTTGGTGGTTTTATAGCCATTTCCATCATCATGTTCTCTCTCG 1020  
 QY 119 ----- 119  
 Db 1021 CTGGGGGTTATCTTAGTGCTCTCATGAATCGGGTGTTCATTAATTTCTCTGAGTTTC 1080  
 QY 119 ----- 119  
 Db 1081 CTTGTGCACTGCGCGTTGGACCTTTGAGTGTGATGCTTTTTCACCTTCTTCACAT 1140  
 QY 119 ----- 119  
 Db 1141 TCTCATGCAAGTCACCCACCATATGTCATAGCCATGAAAGCAAGCAATGAAAAAGA 1200  
 QY 119 ----- 119  
 Db 1201 GGAACACTTTTCAGTCATCTGTCTTCTCAAAAATAGAAAGAAAGTCTATTTTGATTC 1260  
 QY 119 ----- 119  
 Db 1261 ACGTGAAGGCTTAACAAGCTCTAGAGGCGCTGATTTCAAGTTTCTTGTGAACATGTC 1320  
 QY 119 ----- 119  
 Db 1321 CTCACATGATCAAACAATTAAAGATAGAAAGAAAAAATCAGAGAAACCTGAAAT 1380  
 QY 119 ----- 119  
 Db 1381 GATGATGATGTGAGATTAAAGAGAGTGTCTCAAGTATGAATCTCACTTTCAACAAT 1440  
 QY 119 ----- 119  
 Db 1441 GAGGAGAAAGTATGATACAGATGATCAACTGAAGGCTATTTACGACAGACTCACAGAG 1500

QY 119 ----- 119  
 Db 1501 CCTCCACCTTGTATCTCAGACGCTGAGCTTGTGAAGAAAGAGTCATGATAGCT 1560  
 QY 119 ----- 119  
 Db 1561 CATGCTATCCACAGAAAGTCTACATGAATATGTACCAGAGGCTGCAGAAATTAATGC 1620  
 QY 119 ----- 119  
 Db 1621 CATTCACATTTCCACGATACACTGGCCAGTCAGACATCTCATTCACCAATCATGAC 1680  
 QY 120 ----- 120  
 Db 1681 TACCATCATATTTCCATCATCATCACACCAACCAAAACCACTCTCACAGTCACAGCAG 1740  
 QY 128 ArgTysSerArgGluGluLeuLysAspAlaGlyValAlaThrLeuAlaTrpMetValIle 147  
 Db 1741 CGCTACTCTGGGAGGAGCTGAAGATGCCGGCTCCCACTTGGCTGATGTGATATA 1800  
 QY 148 MetGlyAspGlyLeuHisAsnPheSerAspGlyLeuAlaIleGlyAlaAlaPheThrGlu 167  
 Db 1801 ATGGGTATGGCGCTGCACAAATTTCAAGCATGGCTTACCAATGGCTGCTTTTACTGAA 1860  
 QY 168 GlyLeuSerSerGlyLeuSerThrSerValAlaValPheCysHisGlyLeuProHisGlu 187  
 Db 1861 GGCTTATCAAGTGTATTAAGTACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1920  
 QY 188 LeuGlyAspPheAlaValLeuLeuLysAlaGlyMetThrValLysGlnAlaValLeuTyr 207  
 Db 1921 TTAGTGACTTTGTGTGTCTTCTTAAGAGCTGGCATGACCGTTAAGCAGGCTGCTTTAT 1980  
 QY 208 AsnAlaLeuSerAlaMetLeuAlaIleTyrLeuGlyMetAlaThrGlyIlePheIleGlyHis 227  
 Db 1981 AATGCAATTCAGCAGCAGCTGGCGCTATCTTGAAATGCAACAGAAATTTTCATATGTCAT 2040  
 QY 228 TyrAlaGluAsnValSerMetTrpIlePheAlaLeuThrAlaGlyLeuPheMetHisVal 247  
 Db 2041 TATCTGAATAATGTTTCTATAGTGAATTTTGACACTTACTGCTGGCTTATTCATATGATGTT 2100  
 QY 248 AlaLeuValAspMetValProGluMetLeuHisAsnAspAlaSerAspHisGlyCysSer 267  
 Db 2101 GCTCTGGTGTATAGTACCTGAAGTCTGCAATGATGATGATGATGATGATGATGATGATGATGAT 2160  
 QY 268 ArgTrpGlyTyrPhePheLeuGlnAsnAlaGlyMetLeuLeuGlyPheGlyIleMetLeu 287  
 Db 2161 CGCTGGGGGTAATTTCTTTTACAGAAATGCTGGATGCTTTTGGGTTTGGAAATATATGTTA 2220  
 QY 288 LeuIleSerIlePheGluHisLysIleValPheArgIleAsnPhe 302  
 Db 2221 CTTATTTCCATATTTGAAACATTAATAATCGTGTTCGTATTAATTTTC 2265  
 RESULT 15  
 ADL57160  
 ID ADL57160 standard; DNA; 2268 BP.  
 AC ADL57160;  
 DE 03-JUN-2004 (first entry)  
 XX Human NOV9d gene seq ID NO:105.  
 KW ds; gene; human; antidiabetic; anorectic; cardiant; hypotensive;  
 KW antiarteriosclerotic; anorectic; viruside; antibacterial; fungicide;  
 KW procoagide; nootropic; neuroprotective; antiparkinsonian;  
 KW anticonvulsant; osteopathic; antiarthritic; antiinflammatory;  
 KW dermatological; antisclerotic; antidiabetic; gene therapy;  
 KW fibroblast growth factor receptor 4; FGF4;  
 KW complement factor I precursor; matrix metalloproteinase-15 precursor;  
 KW MD3; T-lymphocyte surface antigen ly-9 precursor;  
 KW fibroblast growth factor-21; FGF-21;  
 KW alpha-2 macroglobulin-like polypeptide variant;

KW antileukoprotease 1 precursor; LIV-1; nuclear hormone receptor NOR-1;  
 KW transmembrane protein-like; beta-neoendorphin-dynorphin precursor.  
 OS Homo sapiens.  
 XX  
 XX  
 XX Key Location/Qualifiers  
 FT CDS 1..2268  
 FT /tag= a  
 XX  
 XX WO2004022723-A2.  
 XX  
 XX 18-MAR-2004.  
 XX  
 XX 09-SEP-2003; 2003WO-US028141.  
 XX  
 XX 09-SEP-2002; 2002US-0409145P.  
 XX 10-SEP-2002; 2002US-0409544P.  
 XX 12-SEP-2002; 2002US-0410320P.  
 XX 16-SEP-2002; 2002US-0411060P.  
 XX 23-SEP-2002; 2002US-0412766P.  
 XX 23-SEP-2002; 2002US-0412825P.  
 XX 24-SEP-2002; 2002US-0412767P.  
 XX 25-SEP-2002; 2002US-0413342P.  
 XX 30-SEP-2002; 2002US-0414832P.  
 XX  
 XX (CURA-) CURAGEN CORP.  
 PA  
 P1 Zhong M, Guo X, Anderson DW, Ort T, Padigaru M, Rieger DK;  
 XX  
 XX WPI: 2004-315567/29.  
 DR P-PSDB; ADL57161.  
 XX  
 FT New isolated NOVX polypeptides and polynucleotides, useful for  
 PT preventing, diagnosing or treating NOVX-associated disorders, e.g.  
 PT osteoarthritis, obesity, arteriosclerosis, cancer, Parkinson's disease,  
 PT asthma, or infections.  
 XX  
 PS Claim 17; SEQ ID NO 105; 214pp; English.  
 XX  
 XX The invention relates to a novel isolated polypeptide (NOVX) comprising a  
 CC mature form of any of the 37 amino acid sequences fully defined in the  
 CC specification. A polypeptide of the invention has antidiabetic,  
 CC anorectic, cardiant, hypotensive, antihypertensive, anorectic,  
 CC vitruce, antibacterial, fungicide, protozoacide, nootropic,  
 CC neuroprotective, antiparkinsonian, anticonvulsant, osteopathic,  
 CC antidiabetic, antiinflammatory, dermatological, antiallergic, and  
 CC antipneumatic activity. A polynucleotide of the invention may have a use  
 CC in gene therapy. The polypeptides, nucleic acid molecules and antibodies  
 CC are useful in the manufacture of a medicament for treating a syndrome  
 CC associated with a human disease, preferably a NOVX-associated disorder.  
 CC The nucleic acid molecules, polypeptides and antibodies are useful for  
 CC treating, preventing or diagnosing diseases such as metabolic disorders,  
 CC diabetes, obesity, infectious diseases (viral, bacterial, fungal,  
 CC helminthic, and protozoal), anorexia, cancer, cardiovascular diseases  
 CC (hypertension, atherosclerosis), neurodegenerative disorders, Alzheimer's  
 CC disease, Parkinson's disease, epilepsy, immune disorders  
 CC (osteoarthritis), haematopoietic disorders, inflammatory skin disorders,  
 CC asthma, and various dyslipidaemias. The nucleic acids and polypeptides  
 CC may also be used as targets for the identification of small molecules  
 CC that modulate or inhibit e.g. neurogenesis, cell differentiation, cell  
 CC proliferation, haematopoiesis, wound healing and angiogenesis, in gene  
 CC therapy, in generation of antibodies that bind immunospecifically to NOVX.  
 CC Substances for use in therapeutic or diagnostic methods. The nucleic  
 CC acids are further used as hybridisation probes, in chromosome mapping,  
 CC tissue typing, preventive medicine, and pharmacogenomics. The NOVX  
 CC polypeptides of the invention show homology to certain known human  
 CC proteins: NOV9a-1c show homology to fibroblast growth factor receptor 4  
 CC (FGFR4); NOV9a shows homology to complement factor 1 precursor; NOV3a  
 CC shows homology to matrix metalloproteinase-15 precursor; NOV4a shows  
 CC homology to MDC3; NOV5a-5c show homology to T-lymphocyte surface antigen  
 CC Ly-9 precursor; NOV6a-6m show homology to fibroblast growth factor-21  
 CC (FGF-21); NOV7a-7c show homology to alpha-2 macroglobulin-like  
 CC polypeptide variant; NOV8a-8g show homology to antileukoprotease 1

CC precursor; NOV9a-9i show homology to LIV-1 protein; NOV10a shows homology  
 CC to nuclear hormone receptor NOR-1; NOV11a-11j show homology to  
 CC transmembrane protein-like; NOV12a-12c show homology to beta-neoendorphin  
 CC -dynorphin precursor. The present sequence encodes a NOVX polypeptide of  
 CC the invention.  
 XX  
 XX  
 XX SQ Sequence 2268 BP; 661 A; 512 C; 476 G; 619 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 5,01e-129 Length: 2268  
 Score: 1356.50 Matches: 301  
 Percent Similarity: 40.00% Conservative: 1  
 Best Local Similarity: 39.87% Mismatches: 0  
 Query Match: 84.83% Indels: 453  
 DB: 12 Gaps: 1  
 US-10-659-004-104 (1-302) x ADL57160 (1-2268)  
 QY 1 MetAlaArgLysLeuSerValIleLeuIleLeuThrPheAlaLeuSerValThrAspPro 20  
 DB 1 ATGGGAGAGAACTTATCTGTAATCTTGATCTTGACCTTGGCCCTCTGTGACAAACCC 60  
 QY 21 LeuHISGluLeuLysAlaAlaAlaPheProGlnThrGluLysIleSerProAsnTyr 40  
 DB 61 CTTGATGAACCTAAAGACAGCTGCTTCCCGACAGCACTGAGAAATTAATCGCAATTGG 120  
 QY 41 GluSerGlyIleAsnValAspLeuAlaIleSerThrArgGlnTyrHisLeuGlnLeu 60  
 DB 121 GAATCGGCACTTAATGTTGATCTTGACCTTGCAATTTCCACGCGCAATATCAACACGCTT 180  
 QY 61 PheTyrArgTyrGlyLysAsnAsnSerLeuSerValGluGlyPheArgLysLeuGln 80  
 DB 181 TTCTACCGCTATGGAGAAATTAATTTCTTGACGTGGGGGATTCAGAAATTAATCTTCA 240  
 QY 81 AsnIleGlyIleAspLysIleLysArgIleHisIleHisIleAspHisAspHisSer 100  
 DB 241 AATATAGCATATGATATGATTAATAAGATCATATATACACGACGACCATCATCA 300  
 QY 101 AspHisGluHisHisSerAspHisGluArgHisSerAspHisGluHisHisSerAsp 119  
 DB 301 GACCACGACATCACTACAGACCATGAGCGGTCACTACAGCATCACTACAGACGAC 360  
 QY 119 ----- 119  
 DB 361 GAGCATCACTGACCATGATCATCATCTGACCAATATCATGCTGCTTGTGTAATAAT 420  
 QY 119 ----- 119  
 DB 421 AAGGAAAAGCTCTTTGCCGACCATGATCACTCAATAGTTCAGGTAAGATCTTAAGAAC 480  
 QY 119 ----- 119  
 DB 481 AGCCAGGGAGAAAGAGCTCACCGACCAAGACATGCGAGGTGAGAGGAATGTCAGAGAC 540  
 QY 119 ----- 119  
 DB 541 AGGTGATGCTAGTGAAGTGAAGTCACTCACTGTGTAACAAGTCTGTAAGAACTCAC 600  
 QY 119 ----- 119  
 DB 601 TTTCTAGAGACATATGAGACTTCAGACCTGGAAAACTTTCCCAAGATGTAGACAGC 660  
 QY 119 ----- 119  
 DB 661 TCACATCCACCATGATGATCATCAAGAGCCGGGTGAGCCGGTGTGTGTAAGAAAACA 720  
 QY 119 ----- 119  
 DB 721 AATGAATGTGTAGTGAAGCCCGAAAAGCTTATGTATTCAGAAAACAATGAAT 780  
 QY 119 ----- 119  
 DB 781 CCTGAGAGTGTTCATATGATCAAGAGTACTGATCATGCGATGGGATTCAGAGTT 840



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

# OM protein - nucleic search, using frame\_plus2n model

Run on: March 19, 2005, 22:48:51 ; Search time 3692 Seconds  
(without alignments)  
3113.600 Million cell updates/sec

Title: US-10-659-004-104  
Perfect score: 1599  
Sequence: 1 MARKSLITLITLALSVTNP.....FGIMLLISIFEHKIVFRINF 302

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues  
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame+2n,model -DEV=xlh  
-Q/cgnt2 1/USPRO.epool/US10659004/funat.14032005.181235.15755/app.query.fasta\_1.455  
-DB=EST -QMT=fastap -SUFFIX=ret -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blonum62 -TRANS=human40.cdi -LIST=45  
-LOCALIGN=200 -THR\_SCORE=pcr -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10659004.OCGN.1.1.3437.0funat.14032005.181235.15755 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEROBJRY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -MAIN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:  
1: gb\_eest1.\*  
2: gb\_eest2.\*  
3: gb\_hcc.\*  
4: gb\_eest3.\*  
5: gb\_eest4.\*  
6: gb\_eest5.\*  
7: gb\_eest6.\*  
8: gb\_gsest1.\*  
9: gb\_gsest2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	1237.5	77.4	3834 3	AK028976 Mus muscu
2	1007	63.0	859 3	CA488910 AGENCOURT
3	981	61.4	864 6	CD251026 AGENCOURT
4	975	61.0	756 6	CA327621 UT-M-FRO-
5	974	60.9	906 6	BU541953 AGENCOURT
6	969	60.6	605 4	B1284823 UI-R-CX08
7	948	59.3	665 7	CF909638 A0539C05-
8	935	58.5	875 5	BX327149 BX327149
9	906	56.7	893 6	CA454784 AGENCOURT

10	893	55.8	582 5	BP259977
11	880	55.0	548 5	BU479906
12	875	54.7	791 4	BG547794
13	850	53.2	539 6	CB052402
14	850	53.2	639 1	BU614923 UI-M-FRO-
15	846	52.9	743 1	A1907176 RC-BT133-
16	838	52.4	746 4	BG530601 BG5260191
17	836.5	52.3	571 2	BH373870 601227885
18	812.5	50.8	1184 3	CR647282 Tetracodon
19	812	50.8	1216 3	CR646555 Tetracodon
20	806	50.4	1049 3	CR639910 Tetracodon
21	803.5	50.3	1134 3	CR636655 Tetracodon
22	801	50.1	1134 3	BF032013
23	777	48.6	1059 4	BM557996
24	772	48.3	692 7	CF147123
25	771	48.2	691 7	CF146933
26	759.5	47.5	1135 4	BM560789
27	753	47.1	980 4	BG994120 PMO-HT091
28	740	46.3	950 5	BQ433766
29	739	46.2	699 7	CG641564
30	738	46.2	486 2	BP913546
31	736	46.0	855 5	BU159290
32	728	45.5	849 2	BF680209
33	719	45.0	766 5	BM948026
34	717.5	44.9	640 7	CO354780
35	717	44.8	757 6	CD298707
36	713	44.6	5235 3	CR749813
37	712	44.5	3706 3	AK049099
38	711	44.4	480 2	BF913542
39	710	44.4	480 2	BF913542
40	707	44.2	3924 3	AK030685
41	701	43.8	580 2	BP915495
42	696	43.5	536 2	BF915454
43	695.5	43.5	751 7	CN526184
44	695.5	43.5	872 1	AU120027
45	688	43.0	632 5	BX925865

## ALIGNMENTS

RESULT 1	AK028976	LOCUS	3834 bp	mus musculus 10 days neonate skin cDNA, RIKEN full-length enriched library, clone:4732477B07 product:ESTROGEN REGULATED LIV-1 PROTEIN homolog (Homo sapiens), full insert sequence.
ACCESSION	AK028976	VERSION	AK028976.1	GI:26324937
KEYWORDS	HTC; CAP trapper.	SOURCE	Mus musculus (house mouse)	
ORGANISM	Mus musculus	REFERENCE	1	Carninci, P. and Hayashizaki, Y.
AUTHORS	Itch, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.	TITLE	High-efficiency full-length cDNA cloning	
JOURNAL	Meeth. Enzymol. 303, 19-44 (1999)	MEDLINE	99279253	
PUBMED	10349636	REFERENCE	2	Carninci, P. and Hayashizaki, Y.
AUTHORS	Itch, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.	TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes	
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)	MEDLINE	20499374	
PUBMED	11042159	REFERENCE	3	Shibata, K., Itch, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kiteunai, T., Tashiro, H., Itch, M., Smi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kaishiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanishi, M.,

TITLE	JOURNAL	REFERENCE	AUTHORS
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuzaki, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.	RIKEN integrated sequence analysis (RISA) system--384-format	Genome Res. 10 (11), 1757-1771 (2000)	205309613
Sequencing pipeline with 384 multicapillary sequencer			11076891
4			
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium	Functional annotation of a full-length mouse cDNA collection	Nature 409, 685-690 (2001)	5
5			
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	Nature 420, 563-573 (2002)	6 (bases 1 to 3834)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hasegawa, T., Hara, A., Hashizume, M., Hayashida, K., Hayatsu, N., Hizumoto, K., Hirooka, T., Hirozane, T., Horii, F., Imocani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishii, K., Nomura, K., Numazaki, R., Ohno, M., Ohtsuo, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sugabe, Y., Tagami, M., Tagawa, A., Takehashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.	Submitted (15-JUL-2001) Yoshinide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.c.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)		
CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.	Please visit our web site for further details.	URL: http://genome.gsc.riken.jp/	
Location/Qualifiers			
1..3834			
/organism="Mus musculus"			
/mol_type="mRNA"			
/strain="C57BL/6J"			
/db_xref="FANTOM_DB:4732477B07"			
/db_xref="taxon:10090"			
/clone="4732477B07"			
/tissue_type="Skin"			
/clone_lib="RIKEN full-length enriched mouse cDNA library"			
/dev_stage="10 days neonate"			
500..2797			
/note="unnamed protein product; ESTROGEN REGULATED LIV-1 PROTEIN homolog [Homo sapiens] (SPR1013433, evidence: FASTY, 86.3%id, 100%length, match=2295)			
putative"			
/codon_start=1			
/protein_id="BAC26223.1"			
/db_xref="GI:26324938"			
/translation="MATLVSVGMILTFALWTSPLHQLQSTAAFSQTEKINSNWEPCVNVDAIVTMORHNLQATLFYVYGVNDLSVEGFRKLQNIIGDIKIRVNIHDEHHAALHEHSHDHEHSHDHEHSHDHEHSHDHSRSHVAVGNKRAFCEDLSDSDHSGKNDRTSLGKSRPAHNNGRNINKSASSEVYAVNVAVSETRVETIEPKPPTRTYCNSTPSITKRSRYGLSLAKKSNEVSECPKRSYMYRNNNDIOCPNTKLTSHEMSIOALNATFNFLCPAIIINQIDARRCAIHTYSEKAAIIPPTYSQITWLGEGTAISISFLNGLVPLNNRVFPLNVAALVAGTISGALHLLPHSHS			

polya_signal		3819..3824	/note="putative"
polya_site		3834	/note="putative"
ORIGIN		/note="putative"	
Alignment Scores:			
Pred. No.:	7,68e-125	Length:	3834
Score:	1237.50	Matches:	276
Percent Similarity:	38.43%	Conservative:	18
Best Local Similarity:	36.08%	Mismatches:	8
Query Match:	77.39%	Indels:	463
DB:	3	Gaps:	2
US-10-659-004-104 (1-302) x AK028976 (1-3834)			
Qy	1	MetaIaArglybLeuSerValIleLeuIleuThrPheAlaLeuSerValThrAsnPro	20
Db	500	ATGACCCACAGATTATATCTGTATATGATCTTGACCTTGCCCTTTGGGTTTACAAAGCCC	559
Qy	21	LeuHISgluLeuIyala--AlaIaPheProGlnThrIrgIuLySileSerProAsn	39
Db	560	CTTCATGACTACATCAATCAACAGCTGCTTTCTCTGAGACTACTGAGAAATTAATCAAT	619
Qy	40	TrpGluSerGlyIleAsnValAspLeuAlaIleSerThrArgGlnTyHisLeuGlnIln	59
Db	620	TGGGAACCTGGTGTAAATGTTGACTTGGCACTTACCATGACGACCAACCATCTCGACGAG	679
Qy	60	LeuPheTyraArgTyrgIygluAsnAsnSerLeuSerValGluGlyPheArgLybLeuIeu	79
Db	680	CTATTTCTACCGCTACGGAGAGAAATGATTCCTTGTCTGTGAAGCTTCGAAAAATTGCTT	739
Qy	80	GlnAsnIleGlyIleAspLySileLyAsnArgIleHisIleHisIAspHisAspHisIb	99
Db	740	CAGAACAATAGGCATATGATTAAGATTAAAAAGATCCATATACACCATGACCAAGCATCAT	799
Qy	100	SerAspHisgluHisIleSerAspHisgluArgHisSerAspHisgluHisIleSerAsp	119
Db	800	GCTGACCAAGCATCACTCGGACCATGAGCATCACTCGGACCAAGCATCACTCGGAC	859
Qy	119	-----	119
Db	860	CACAGCATCACTCGGACCAAGCATCACTCGGACCAAGCATCACTCGGACCATTCACCGCAGT	919
Qy	119	-----	119
Db	920	CACAGGTTGCTGGTAAAAACAATCGGAAAGCCTTTTGTCCAGACTTGACTGTATAT	979
Qy	119	-----	119
Db	980	TCAGTAAAAATCCTAGAACTAGTCTAGGAAAGGATCTCGCCAGACAGACATGAT	1039
Qy	119	-----	119
Db	1040	GGTAGAGAGAAACATCAAGAGAGATGCAAAGCTCTTAGTAAATGATCCTCGGCGGTATACAC	1099
Qy	119	-----	119
Db	1100	GCTGTCTCTGAAGAACTCGCTTTGTAGAGACATATAGAGACTCCAAAACTCGGAGAGGC	1159
Qy	119	-----	119
Db	1160	ACCAAGATGTAAACCTTTCTACCCCAACGACATCACGAGAAAAAGCCGAGTGGCGCG	1219
Qy	119	-----	119

```

Db 1220 CTGAGTCGGCTAGCTAGGAGAGAAAAGCAATGAGTCTGTGAGTGAGCCAGAGAAAGCTTT 1279
Qy 119 ----- 119
Db 1280 ATGTATTCCAGAAAACAAATAGACAATATTGAGAGTGTTCATATACAAACAAGCTGTG 1339
Qy 119 ----- 119
Db 1340 ACATCCCATGGAGATGAGCATCCAGGCTGTGTTGAATGCACGGAATTTAACTATCTCTGC 1399
Qy 119 ----- 119
Db 1400 CCAGCATCATCAATCAAAATTGATGCTGGGCTGTGTGATTCAATACAGCAAGTAGAAG 1459
Qy 119 ----- 119
Db 1460 AAGCAGAAATCCCTCCAAAGACCTATTCTTACAAATAGCCTGGCTTGTGGCTTCATA 1519
Qy 119 ----- 119
Db 1520 GCCATTTCCATCATCAGTTTCTGTCTGTGGAGATCATCTTGGTCCACTCATGAAC 1579
Qy 119 ----- 119
Db 1580 CGGGATATTTTCAAGTTCCTGTAGCTTCTCGTGGCGCTGGCGGTGGAACGCTAGT 1639
Qy 119 ----- 119
Db 1640 GGCAGTCTCTGTACATCTTCTCCACACTCTCATGAGTATCAGCAGATCATAGC 1699
Qy 119 ----- 119
Db 1700 CATGAGAGCCAGGAGTAGAAATGAAAAGAGCCCCCTGTTCAGCCACTGTCCGCTCAG 1759
Qy 119 ----- 119
Db 1760 AATATAGAGAAAGTCTATTATTTTGATTCAGTGTGAAGGTCTGACGGCTCTAGGGGGC 1819
Qy 119 ----- 119
Db 1820 TTATATTTTCATGTTTCTTGTGAAACAGTACTCAGACTGATCAAGCAATTTAAAGATAG 1879
Qy 119 ----- 119
Db 1880 AAAAAGAGAAATCAAAAAGAACTGAAAATGATGAGATGTGAGAGAGAGAGAGAGAGCTG 1939
Qy 119 ----- 119
Db 1940 TCCAAATACGACTCTCAGCTTTCCTCAATAGAGAGAGTGGAGCCAGGGGAACGACT 1999
Qy 119 ----- 119
Db 2000 GAAAGTATCTGCGAGCCGAGCTCCCAAGAGCCCTCCCTTTGATTCCAGAGCCGAGC 2059
Qy 119 ----- 119
Db 2060 ATGTGAGAGAGAGAGAGTATGATGAGCCATGACACCAAGAAAGTCTAATGAA 2119
Qy 119 ----- 119
Db 2120 TATGTGCCAGGGGCTGCAAGAACAGTCCATTCACTTCCAGATACGCTGGCCAG 2179
Qy 119 ----- 119
Db 2180 TCCGAGCACTCATCCACCAACATCAGACTACATTTCTGCAACCAACACACAC 2239
Qy 120 ----- 120
Db 2240 CAGAACCCACCCCTCAGAGCCACGCGGCTATCTTCAGAGAGAGCTGAAAGAGCC 2299
Qy 138 G1yVala1aThrLeu1a1aTrpMetVal11IleMetG1yAerG1yLeu11AspPhseSer 157
Db 2300 GGCATTCGCAATTCGCTGATGATGATCATGGCGAGCGGGCTGCAATTTCACTGAC 2359

```

```

Qy 158 G1yLeu1a1aIleG1yA1aA1aPheThrG1yG1yLeuSerSerG1yLeuSerThSerVal 177
Db 2260 GGCCTGTGATTTGGTGTGCTGCTTCCAGAGGTTTGTCCAGTGGAGCTTAAGACCTCTGTG 2419
Qy 178 A1aVal1PheCyH1eG1yLeuProH1eG1yLeuG1yAerPheA1aVal1LeuLeuYsa1a 197
Db 2420 GCTGTGTTCTGTCAATGAACCTGCTCATGAACTAAGTGAACCTTCTGCTTTTCTTAAAGCT 2479
Qy 198 G1yMetThrVal1yG1yG1a1aVal1Leu1yTrna1a1eUSeSerA1aMetLeuA1aTyrieu 217
Db 2480 GGCATGACTGTCAGACAGGCTGTGCTCTATATGCTCTGTGACGCAATGTGGCTACCTT 2539
Qy 218 G1yMetA1aThrG1y11ePhe11eG1yH1eTrp1yTrp1yA1aUanVal1SerMetTrp11ePhe 237
Db 2540 GGAATGGCAACAGGAGATATTCATCGGCATTATGACAGAAATGTTTCTATGTGATATTC 2599
Qy 238 A1aLeuThrA1aG1yLeuPheMetC1eValA1aLeuVal1aerMetVal1ProG1yMetLeu 257
Db 2600 GCACTCACTGCGGCTTGTTCATATGATGTGCTGTGTGACATGTGCTCGAGATGTG 2659
Qy 258 H1saValPheA1aSerAerPhe1eG1yCySerAerTrpG1yTyrrPhePheLeuG1aVala 277
Db 2660 CACATGATGCTAGTATGATCAGGATGACAGCCGTTGGGATATTTCTTCTCAGATGCT 2719
Qy 278 G1yMetLeuLeuG1yPheG1y11eMetLeuLeu11eSer11ePheG1yH1eYsa1a 297
Db 2720 GGGATACTTTCGGTGTGTGGAAATGATGTACTCATTTCCATATTTGAGCATAAATGTG 2779
Qy 298 PheArg11eAerPhe 302
Db 2780 TTTCTATTAATTTTC 2794

RESULT 2
CA488910 859 bp mRNA linear EST 14-NOV-2002
LOCUS AGENCOURT_10808247 MAPCL Homo sapiens cDNA clone IMAGE:672115 5'
DEFINITION mRNA sequence.
CA488910
VERSION CA488910.1 GI:24951701
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Bukatyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/.
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Kristi A. Eglund, Ira Pastan
CDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://Image.lnl.gov
Plate: L14M14281 row: 1 column: 11
High quality sequence stop: 755.
FEATURES
location/qualifiers
1..859
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:672115"
/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231,
hTERT-IME1, LNCaP"
/lab_host="EMDH108"
/clone_id="MAPCL"
/notes="Vector: PCMV-SPORT6; Site 1: EcoRV; Site 2: Not I;
Subtracted with brain, liver, lung, kidney and muscle.
Directionally cloned. Priming method: oligo-dT. Average

```



QY 224 PheleGlyHisETyAlaGluValSerMetTrpIlePheAlaLeuThrAlaGlyLeu 243  
 DB 512 TTCATTGGTCATTATGCTGAAAATGTTTATATGATATTTGGCACTTATCTCTGCTTA 571  
 QY 244 PheMetHisValAlaLeuValAspMetValProGluMetLeuHisAsnAspAlaSerAsp 263  
 DB 572 TTCATGATATGTTGCTCTGTTGATATGTAACCTGAAATCTGCACATGATCTATGATGC 631  
 QY 264 HisGlyCysSerArgTrpGlyTyrrPhePheLeuGlnAsnAlaGlyMetLeuGlyPhe 283  
 DB 632 CATGATATGAGCGCTGGGGGTATTTCTTTTACAGAACTGGGATGCTTTTGGGTTT 691  
 QY 284 -GlyIleMetLeuLeuIleSerIlePheGluHisIleVal-PheArgIleAsnPro 302  
 DB 692 GGGATTATGTTACTTATTTCCATTTTCAACATTAATAATGCTGTTTGGTAAATTTTC 750  
 RESULT 4 756 bp mRNA linear EST 09-JUL-2003  
 LOCUS CA327621  
 DEFINITION U1-M-F02-j-07-0-UI.r1 NIH\_BMAP\_F10 Mus musculus cDNA clone  
 CA327621  
 IMAGE:6825568 5', mRNA sequence.  
 ACCESSION CA327621.1 GI:24545719  
 VERSION EST.  
 KEYWORDS Mus musculus (house mouse)  
 SOURCE  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.  
 NIH-MGC http://mgi.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs-remail.nih.gov  
 Tissue Procurement: Dr. Jim Lahn, University of Iowa  
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
 DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/BLN at:  
 http://image.llnl.gov  
 This clone was contributed by the Brain Molecular Anatomy Project  
 (BMAP)  
 The following repetitive elements were found in this cDNA  
 sequence: 63-144, >(TGG)nSimple\_repeat (matched complement)  
 Seq primer: PYX-5.  
 Location/Qualifiers  
 1..756  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="CS7BL/6"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:6825568"  
 /tissue\_type="whole brain"  
 /dev\_stage="embryo 13.5,14.5,16.5,17.5dpc"  
 /lab\_host="DH10B (T1 phage resistant)"  
 /clone\_lib="NIH\_BMAP\_FY0"  
 /note="Organ: Brain; Vector: pYX-Asc; Site 1: Ecor I;  
 Site 2: Not I; The library was constructed according  
 to Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
 1996. Denatured RNA was size fractionated on a 1% agarose  
 gel. First strand cDNA synthesis was primed with oligo-dT  
 primer containing a Not I site. Double strand cDNA was  
 size selected according to mRNA size fraction, ligated  
 with Ecor I adaptor, digested with NotI and then cloned  
 directionally into pYX-Asc vector. The library tag  
 sequence located between the Not I site and the polyA tail  
 is AGCGACAG. This library was created for the University  
 of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the  
 Developing Mouse Nervous System', supported by National  
 Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,  
 program coordinator."

ORIGIN

Alignment Scores:  
 Pred. No.: 6,87e-97 Length: 756  
 Score: 975.00 Matches: 190  
 Percent Similarity: 92.42% Conservative: 5  
 Best Local Similarity: 90.05% Mismatches: 6  
 Query Match: 60.98% Indels: 11  
 DB: 6 Gaps: 2  
 US-10-659-004-104 (1-302) x CA327621 (1-756)  
 QY 92 ILeHisAspArgHisAspHisHisSerAspHisGluHisSerAspHisGluArgHis 111  
 DB 61 ATCCACCCACATCACGATACCAT-----CACATTCTGCAC 96  
 QY 112 SerAspHisGluHisHisSerAspHisHisProHisSerHisSerGlnArgTyrrSerArg 131  
 DB 97 -----CACACACACACACACACACACACCTCAGACACACACACACGCTACTCTCGA 150  
 QY 132 GluGluLeuValAspAlaGlyValAlaThrIleuAlaTrpMetValIleMetGlyAspGly 151  
 DB 151 GAGGAGCTGAAGAGACGCGGATTCACATTTGGCTGGATGTCATATGCGGACGCG 210  
 QY 152 LeuHisAspHisSerAspGlyLeuAlaTrpGlyAlaAlaPheThrGluGlyLeuSerSer 171  
 DB 211 CTGCACAAATTCAGTACGCGGCTTCTATTTGGTCTCTTCCACGAGGTTTGTCCAGT 270  
 QY 172 GlyLeuSerThrSerValAlaValPheCysHisGluLeuProHisGluLeuGlyAspPhe 191  
 DB 271 GGGCTTAACACCTCTGTGCTGTGTCTGTATGATGAACGCTCAGTAAGTAACTAGTACTTT 330  
 QY 192 AlaValLeuLeuValAlaGlyMetThrValIleGlnAlaValLeuTyrrAsnAlaLeuSer 211  
 DB 331 GCTGTTTTCGTAAGAGCTGGCATGATGCTGCAAGCAGGCTGCTCTATTAATGCTCTGCA 390  
 QY 212 AlaMetLeuAlaTyrrLeuGlyMetAlaThrGlyIlePheIleGlyHisTyrrAlaGluAsn 231  
 DB 391 GCCATGTTGGCTTACCTTGGAAATGGACACAGAGATATTCATGAGGATTAATGCAAAAT 450  
 QY 232 ValSerMetTrpIlePheAlaLeuThrAlaGlyLeuPheMetHisValAlaLeuValAsp 251  
 DB 451 GTTCTATGATGATATTCGCACTACCTACCTCGGCTTTGTCATGATATGCTCTGTTAC 510  
 QY 252 MetValProGluMetLeuHisAsnAspAlaSerAspHisGlyCysSerArgTrpGlyTyrr 271  
 DB 511 ATGGTGCTGAGATGTCACAAATGATCTGATGTCACGATGACAGCGCTGGGAGATAT 570  
 QY 272 PhePheLeuGlnAsnAlaGlyMetLeuGlyPheGlyIleMetLeuLeuIleSerIle 291  
 DB 571 TTCTTCTGCAAGATGCTGGATATCTTCTCGTTTGGAAATTAATTAATTAATTCATA 630  
 QY 292 PheGluHisIleValAlaPheArgIleAsnPro 302  
 DB 631 TTGAGCATTAATT-GTCTTTCGTAATAATTC 662  
 RESULT 5 906 bp mRNA linear EST 13-SEP-2002  
 LOCUS BUS41953  
 DEFINITION AGNC00RT.10325419 NIH\_MGC\_40 Homo sapiens cDNA clone IMAGE:6573733  
 5', mRNA sequence.  
 ACCESSION BUS41953  
 VERSION BUS41953.1 GI:22852436  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 NIH-MGC http://mgi.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs-remail.nih.gov





Alignment Scores:	
Pred. No.:	2,32e-92
Score:	935.00
Percent Similarity:	89.04%
Best Local Similarity:	87.21%
Query Match:	58.47%
DB:	5
Gaps:	1
Length:	875
Matches:	151
Conservative:	4
Mismatches:	22
Indels:	1
Gaps:	1

Tissue Procurement: Kristi A. Eglund, Ira Pastan  
cDNA Library Preparation: Invitrogen Corp.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
Plate: L1AM14282 row: 1 column: 11  
High quality sequence stop: 593.

FEATURES  
source  
Location/Qualifiers  
1..893  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6721427"  
/cell\_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231,  
hTERT-HME1, LNCaP"  
/lab\_host="EMD108"  
/clone\_lib="MAPCL"  
/note="Vector: pCMV-SPORT6; Site 1: EcoRV; Site 2: Not I;  
Subcloned with brain, liver, lung, kidney and muscle.  
Directionally cloned. Priming method: oligo-dT. Average  
insert size: 1800 bp. Library amplification: 26,000 fold.  
Kristi A. Eglund, James J. Vincent, Robert Strausberg,  
Bunkook Lee & Ira Pastan: Discovery of new breast  
cancer genes encoding membrane and secreted proteins.  
Manuscript submitted."

Alignment Scores:  
Pred. No.: 3.65e-89 Length: 893  
Score: 906.00 Matches: 175  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.43% Mismatches: 0  
Query Match: 56.66% Indels: 0  
DB: 6 Gaps: 0

US-10-659-004-104 (1-302) x CA454784 (1-893)

Qy 127 GlnaGTYrSerARgGluGluLeuYsaApAlaGlyValAlaThrLeuAlaTrpMetVal 146  
Db 2 CAGGCGTACTCTCGGAGAGAGCTGAAGAATGCGGCGTGCACCTCGCCGTGAGTGTG 61

Qy 147 IlleMetGlyAspGlyLeuHlaAsnPheserAspGlyLeuAlaIleGlyAlaAlaPheThr 166  
Db 62 ATAAATGGGTGATGGCTGCACAAATTTCAAGCATGGCCATGCAATGGCTGCTTACT 121

Qy 167 GluGlyLeuSerSerglyLeuSerThrSerValAlaValPheCyHlaGluLeuProHla 186  
Db 122 GAAGCTTATCAAGTGGTTAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTCAT 181

Qy 187 GluLeuGlyAspPheAlaValLeuLeuYsaIaGlyMetThrValGlyAlaValLeu 206  
Db 182 GAATTAAGTGAATTTGCTGCTTACTTAAAGCTGCAAGCCGTTAAAGCAGCTGCTCT 241

Qy 207 TyrAsnAlaLeuSerAlaMetLeuAlaTyrLeuGlyMetAlaThrGlyIlePheIleGly 226  
Db 242 TATAATGATGTCAGCCATGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 301

Qy 227 HlaTyrAlaGluAsnValSerMetTrpIlePheAlaLeuThrAlaGlyLeuPheMetHla 246  
Db 302 CATATGCTGAAATGTTCTTATGATGATTTGACATTACCTGCTGCTGCTTATTCATGAT 361

Qy 247 ValAlaLeuValAspMetValProGluMetLeuHlaAsnAspAlaSerAspHlaGlyCys 266  
Db 362 GTTGCTGCTGTTGATATGTAATGCTGAATGCTGCAATGATGCTAGGACCATGAGAT 421

Qy 267 SerArgTrpGlyTyrPhePheLeuGluAsnAlaGlyMetLeuLeuGlyPheGlyIleMet 286  
Db 422 AGCCGCTGGGGGATTTCTTTTACAGAAATGCTGGAGATGCTTTGGGTTTGGAAATTATG 481

Qy 287 LeuLeuIleSerIlePheGluHlaIleYsaIleValPheArgIleAsnPhe 302  
Db 482 TTACTATATTCATATTTGAACATTAATAATCGTGTGCTATTAATTTTC 529

RESULT 10  
BP259977  
LOCUS BP259977 Sugano cDNA library, heart Homo sapiens cDNA clone  
DEFINITION HRT08420, mRNA sequence.  
ACCESSION BP259977  
VERSION BP259977.1 GI:52175207

KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,  
Mizushima-Sugano,J., Nakai,K. and Sugano,S.  
Sequence comparison of human and mouse genes reveals a homologous  
block structure in the promoter regions  
Genome Res. 14 (9), 1711-1718 (2004)  
JOURNAL  
COMMENT Contact: Yutaka Suzuki  
Department of Virology  
Institute of Medical Science, University of Tokyo  
4-6-1, Shirokane-dai, Minato-ku, Tokyo 108-8639, Japan  
Email: yusuzuki@ims.u-tokyo.ac.jp.

FEATURES  
source  
Location/Qualifiers  
1..582  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="HRT08420"  
/tissue\_type="heart"  
/clone\_lib="Sugano cDNA library, heart"

Alignment Scores:  
Pred. No.: 5.31e-88 Length: 582  
Score: 893.00 Matches: 177  
Percent Similarity: 99.44% Conservative: 1  
Best Local Similarity: 98.88% Mismatches: 1  
Query Match: 55.85% Indels: 1  
DB: 5 Gaps: 0

US-10-659-004-104 (1-302) x BP259977 (1-582)

Qy 124 SerHisSerGlnaGTYrSerARgGluGluLeuYsaApAlaGlyValAlaThrLeuAla 143  
Db 2 AGTACACAGCCGCTACTCTCGGAGAGAGCTGAAGA-GCCGCGTGCACCTCTGACC 60

Qy 144 TrpMetValIleMetGlyAspGlyLeuHlaAsnPheserAspGlyLeuAlaIleGlyAla 163  
Db 61 TGGATGCTGATTAAGGCTGATGCTGCACAAATTTCAAGCATGGCCATGCAATGGCTGCT 120

Qy 164 AlaPheThrGluGlyLeuSerSerglyLeuSerThrSerValAlaValPheCyHlaGly 183  
Db 121 GCTTTTACTGAAGCTTATCAAGTGGTTAACTTCTGCTGCTGCTGCTGCTGCTGCTGCT 180

Qy 184 LeuProHlaGluLeuGlyAspPheAlaValLeuLeuYsaIaGlyMetThrValIleGln 203  
Db 181 TTGCTCATGAATTAAGGATTTGCTGCTTACTTAAAGCTGCGCATACCGTTAAAGCAG 240

Qy 204 AlaValLeuTyrAsnAlaLeuSerAlaMetLeuAlaTyrLeuGlyMetAlaThrGlyIle 223  
Db 241 GCTGCTCTTAAATGATGCTGACCATGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 300

Qy 224 PheIleGlyHlaTyrAlaGluAsnValSerMetTrpIlePheAlaLeuThrAlaGlyLeu 243  
Db 301 TTCAATGGCTCATTTATGCGAAATGTTCTTATGATGATTTGCACTTACCTGCTGCTT 360

Qy 244 PheMetHlaValAlaLeuValAspMetValProGluMetLeuHlaAsnAspAlaSerAsp 263  
Db 361 TTCATGATATGTTGCTGCTGTTGATGATGCTGAATGCTGCAATGATGCTAGTAC 420

Qy 264 HlaGlyCysSerArgTrpGlyTyrPhePheLeuGluAsnAlaGlyMetLeuLeuGlyPhe 283  
Db 421 CATGATGTAACCGCTGGGGATTTCTTTTACAGAAATGCTGGAGATGCTTTGGGTTTT 480

Qy 284 GlyIleMetLeuLeuIleSerIlePheGluHlaIleYsaIleValPheArgIleAsnPhe 302  
Db 481 GGAATTAATGTAATTTATTCATATTTGAACATTAATAATCGTGTGCTATTAATTTTC 537

RESULT 11



C, or G and N = A, C, G, or T). Average insert size 1.5 kb (range 0.5-4.0 kb). 12/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH\_MGC library."

## ORIGIN

**Alignment Scores:**

Pred. No.:	8.1e-86	Length:	791
Score:	875.00	Matches:	160
Percent Similarity:	99.42%	Conservative:	2
Best Local Similarity:	98.26%	Mismatches:	1
Query Match:	54.72%	Indels:	0
DB:	4	Gaps:	0

QY	131	rrgslugluleuLysAspAlaGlyValaLathrLeuAlaTrpMetValIleMetGlyAsp	150
Db	1	CGGAGAGAGCTGAAGAAGATCCGGCGCTCGCCACTTTGGCTGGATGGTATATGGGTGAT	60
QY	151	GlyLeuH1saenPhseSerAspGlyLeuAlaIleGlyAlaAlaPheThrGluGlyLeuSer	170
Db	61	GGCGTCGACCAATTTCAAGCGATGGCCCTAGCAATTTGGTGCCTTTACTAGAGCTTATCA	120
QY	171	SerGlyLeuSerThrSerValAlaValPheCysH1gLuLeuProH1gLuLeuGlyAsp	190
Db	121	AGTGGTTTAAGACTTCTGTTGCTGCTGTCTGATCAAGATTGCCCTCAATAATTAGGTGAC	180
QY	191	PheH1aValL1leuLeuLysAlaGlyMetThrValLysGlnAlaValLeuLysAsnAlaLeu	210
Db	181	TTTGGTGTTCTACTAAAGCTGGCGATGACCGCTTAAGCAGGCTGTCTCTTATATATGATTG	240
QY	211	SerAlaMetLeuAlaTyrLeuGlyLysMetAlaThrGlyIlePheIleGlyH1strYalGlu	230
Db	241	TCAGCGATGCTGGCGCTATCTTGGAATGGCACAAGGAATTTCAATTCATATATGCTGAA	300
QY	231	AsnValSerMetTrpIlePheAlaLeuThrAlaGlyLeuPheMetH1sValAlaLeuVal	250
Db	301	AAATGTTTCTATGTGATATTTGCCACTTACGTGCGCTTATTCATGATATGTGTCTGGTT	360
QY	251	AspMetValProGlnMetLeuH1saenAspAlaSerAspH1GlyCysSerAlaTrpGly	270
Db	361	GATATGTGACTGAAATGCTGCACATGATGCTAGGACCATGATGATGACCGCTGGGG	420
QY	271	TyrPhePheLeuGlnAsnAlaGlyMetLeuLeuGlyPheGlyIleMetLeuLeuIleSer	290
Db	421	TATATCTTTTAAACAATGCTGGGATGCTTATGGGTTCTCGGAATATATGTTACTATTTC	480
QY	291	IlePheGlnH1sLysIleValPheArgIleAsnPro	302
Db	481	ATATTTGAACATTAATATGCTGTTTGGTATTAATATTC	516

RESULT	13
LOCUS	CB052402
DEFINITION	CB052402 NCSC.G107G06.Y1 NCI_CGAP_Le12 Homo sapiens cDNA IMAGE:3320314
ACCESSION	CB052402
VERSION	5.1, mRNA sequence.
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
TITLE	Eukaryotic Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 539) NCI-CGAP <a href="http://www.ncbi.nlm.nih.gov/ncicgap">http://www.ncbi.nlm.nih.gov/ncicgap</a> . National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index Unpublished (1997) Contact: Robert Strausberg, Ph.D. Email: CGAP@rcmail.nih.gov CDNA Library Preparation:
JOURNAL	
COMMENT	

## FEATURES

**Source**

cDNA library Arrayed by: The I.M.A.G.E. Consortium/LINL DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NIISC)  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: [infoimage.lnl.gov](http://infoimage.lnl.gov)  
Plates: L14M8057 Row: N Column: 11  
Seq primer: M13RPL reverse primer (ABI).

**JRES**  
**Location/Qualifiers**

```

source
1. 539
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:329031.4"
/risuse_type="1elomvosarcoma"
/lab_host="DH10B"
/clone_lib="NCI CGAP Le12"
/notes="Organ: soft tissue; Vector: pTR3D-Pac (Pharmacia)
with a modified polylinker; Site_1: Not I; Site_2: Eco RI.
1st strand cDNA was primed with a Not I - oligo(dT) primer
[5'-AATCGAGAGATTCGGCGCGCATGCTTTTCTTTTCTTTT-3'],
double-stranded cDNA was ligated to Eco RI adaptor
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pTR73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and W. Fatima Bonaldo."

```

## ORIGIN

### Alignment Scores:

Pred. No.:	2.66e-83	Length:	1
Score:	830.00	Matches:	1
Percent Similarity:	100.00%	Conservative:	5
Best Local Similarity:	97.01%	Mismatches:	0
Query Match:	53.16%	Indels:	0
DB:	6	Gaps:	0

US-10-659-004-104 (1-302) X CB052402 (1-539)

QY	130	ASPAIAGIYVAIAIATThLeuAlaATrMetValIILeMeGIYASPIYLeuHISAsnPhe	155
	...	...	
Db	21	GAAGCGGSGTGTCCACTTGTGGCTGGATGGTGAATAGGGTGTGGCTGCACAAATTC	80
QY	156	SeTASPGIYLeuAlaIIeGIYAlaAlaIaphethrIGuGIYLeuSeSeSerGIYLeuSerThr	175
Db	81	AGCGATGCGCTTACGCAATTTGGTCTCTTTACTGAAGCTTATCAAGTGTTTAAGTACT	140
QY	176	SeTVAIAIaValAPheCYAHISGIuLeuProHISGIuLeuGIYASPPheAlaValLeuLeu	195
Db	141	TCGTGTCGTGTCTTGTGCATGAAGTGTGCATGAATTAAGTGTACTTTTGTGTCTACTA	200
QY	196	IYVAIAGIYMeTThValIYSGIAlaValLeuTYrAsnAlaIeUSerAlaMeTLeuAla	215
Db	201	AAGCGTGCATACCGTTAACGACGGCTCTCTTATATATACATTTGTACGCCAAGCTGGCG	260
QY	216	TYrLeuGIYMeTAlaTThGIYrIlePheIleGIYHISrYrAlaGIYAsnValSerMetTrp	235
Db	261	TATCTTGGATGTGCACAGCAATTTTCATTTGCTATTAATCTGAAATGTTCCTAATGTG	320
QY	236	IlePheAlaLeuThAlaGIYLeuPheMetHISValAlaLeuValAspMetValProGlu	255
Db	321	ATATTTTGCACCTTACTGCTGGCTTATTCATGATGTTGTGCTGTGTATAGTACTGAA	380
QY	256	MeTLeuHISAsnAspAlaSerAspHISGIYCYSerArgrTIpGIYrYrPheLeuGln	275
Db	381	ATGCTGCACAAATGATGTACTAGACCAATGATGTACCGCTGGGGGTATTTCTTTTACAG	440
QY	276	AsnAlaGIYMeTLeuLeuGIYPhGeGIYrIleMeTLeuLeuIleSerIlePheGluHISrYs	295
Db	441	AATGCTGGGATGCTTTTGGGTTTGGAAATATGTACTTTTTCATATTTTGAACATAAA	500
QY	296	IleValAPheArGIleAsnPhe	302
Db	501	ATGCGTGTTCGTGTAAATTTTC	521



```

/dev stage="Adult"
/clone.lib="BT133"
/note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No.
196,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

```

## ORIGIN

## Alignment Scores:

Pred. No.:	1,18e-82	Length:	743
Score:	846.00	Matches:	181
Percent Similarity:	86.05%	Conservative:	4
Best Local Similarity:	84.19%	Mismatches:	13
Query Match:	52.91%	Indels:	17
DB:	1	Gaps:	2

US-10-659-004-104 (1-302) x AI907176 (1-743)

```

QY 92 11eh1sh1saeph1saeph1sh1seraeph1sgluh1sh1seraeph1sgluarh1s 111
Db 82 ATTACCCACATCATGACTACCAT-----CATATTCCTCAT 117
QY 112 Seraph1sgluh1sh1seraeph1sh1proh1serh1serg1narq1tyserarg 131
Db 118 -----CATCACACACACCAACACCATCTCACAGTCACAGCCAGCCCTACTCTCG 171
QY 132 GlugluLeuLyAspAlaG1yValAlaThrLeuAlaTpmetVal11emetG1yAspG1y 151
Db 172 GAGAGCTGAAGAAGATGCCGCGCTGCCACTGCGCTGAGATGGTATATATGGATGSC 231
QY 152 LeuH1saeph1seraeph1yLeuAla11leg1yAlaAlaPheThrg1yLeuSerSer 171
Db 232 CTGCACATTTCAAGGATGGCTAGCAATGGTGCTGCTTTACTGAAGGCTTATCAAGT 291
QY 172 GlyLeuSerThrSerValAlaValaPheCysh1sgluLeuProh1sgluLeuG1yAspPhe 191
Db 292 GGTTAAGTACTTCTGTTGCTGTGTTCTGTCATGAGTGGCTCATGAATAGTGACTTT 351
QY 192 AlAval1LeuLeuLyAspAlaG1yMetThrVal1yvg1na1Ala1LeuTyraAspAlaSer 211
Db 352 GCTGTTCTACTTAAGGCTGGCATGACCGTTAAGCAGGCTGTCCTTTATATGCAATTGTCA 411
QY 212 AlAmetLeu-AlaTyrlLeuG1yMetAlaThrG1y11ephe11eg1yH1sTyra1aG1uAs 231
Db 412 GCCATGCTGGGCGTATCTTGGATAGCAACAGGAATTTTCATTGTCATTATGCTGAADA 471
QY 231 nValSerMetTrp11epheAlaLeuThra1aG1yLeuPheMetH1sVala1aLeuValaAs 251
Db 472 TGTTCATATGTGATATTTGCACTTACGCTGCTTATTCATATGTTGCTCTGGGTGA 531
QY 251 pMetVal1ProGluMetLeuH1saAspAla-SerAspH1s-G1yCysSerArg-Trip1 270
Db 532 TAGGNACCTGAAGCTGCAACATGATGCTTAGTACCATGGGATGAGCGGTGGGGG 591
QY 270 YTyrlPhePheLeuGin-AsnAlaG1yMetLeuLeuG1yPhe--G1y11emetLeuLeu11 289
Db 592 GTATTTCTTTTACAAAATGCTGGGATGCTTTNGGGTTTGGGAATTATGTTACTTAT 651
QY 289 eSer11epheGluH1sLy11eValaPheArg 299
Db 652 TTNCATATTGACATMAANCCGGGTTTCA 682

```

Search completed: March 20, 2005, 01:36:15  
 Job time : 3705 secs

**This Page Blank (uspto)**